```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-896 <RES>
A;Cross-references: GB:S79263; NID:g1086954; PIDN:AAB35068.1; PID:g1086955
C;Genetics:
A;Gene: rIL-3Rbeta
C;Superfamily: interleukin-3 receptor beta chain; cytokine receptor homology
C;Keywords: cytokine receptor
F;39-235/Domain: cytokine receptor homology <CRS1>
F;253-433/Domain: cytokine receptor homology <CRS2>

   I56563  Length: 896  July 16, 2003 13:42  Type: P  Check: 5352  ..
Found using 'var2' (kam816.key)

...

                                                          |-----|
421     PEYKGLWSEWSNECTWTTDWVMPTLWIVLILVFLILTFLLALRFGCIYGCKLYRRWKEKI
                                                        471   477

481     PNPSKSLLFQDGGKGLWPPGSTVTFSSKNPTPQGPQNNLFSELQGVS

...

-------
1 match found in sequence:
 s36348 ; TOIG of: s36348  check: 3344  from: 1  to: 234
  (from "var2pir.pep")
  TOIG of: s36348  check: 3344  from: 1  to: 234

 F1;S36348 - opacity protein opa65 - Neisseria gonorrhoeae (strain VP1)
  (fragment)
 N;Alternate names: outer membrane protein opa65
 C;Species: Neisseria gonorrhoeae
 A;Variety: strain VP1
 C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 26-Aug-1999
 C;Accession: S36348; S28617
 R;Kupsch, E.M.; Knepper, B.; Kuroki, T.; Heuer, I.; Meyer, T.F.
 EMBO J. 12, 641-650, 1993
 A;Title: Variable opacity (Opa) outer membrane proteins account for the cell
  tropisms displayed by Neisseria gonorrhoeae for human leukocytes and epithelial
  cells.
 A;Reference number: S36328; MUID:93178439; PMID:8440254
 A;Accession: S36348
 A;Molecule type: DNA
 A;Residues: 1-234 <KUP>
 A;Cross-references: EMBL:Z18940; NID:g49336; PIDN:CAA79373.1; PID:g940802
 A;Experimental source: strain VP1
 A;Note: expression of opacity proteins is regulated by the number of translated
  repeat elements CTCTT, which code for part of the signal sequence; the protein
  can only be synthesized when the number of repeats place the start codon in
  frame with the rest of the protein
 C;Genetics:
 A;Gene: opa65
 C;Superfamily: opacity protein
 C;Keywords: cell surface component; transmembrane protein
 F;1-234/Product: opacity protein opa65 #status predicted <MAT>
 F;11-19/Domain: transmembrane #status predicted <TM1>
 F;20-51/Domain: extracellular #status predicted <EXT1>
 F;28-37/Region: semivariable region
 F;52-60/Domain: transmembrane #status predicted <TM2>
 F;65-71/Domain: transmembrane #status predicted <TM3>
 F;72-107/Domain: extracellular #status predicted <EXT2>
 F;78-107/Region: hypervariable region HV1
 F;108-122/Domain: transmembrane #status predicted <TM4>
 F;128-138/Domain: transmembrane #status predicted <TM5>
 F;139-185/Domain: extracellular #status predicted <EXT3>
 F;144-176/Region: hypervariable region HV2
 F;186-198/Domain: transmembrane #status predicted <TM6>
 F;202-210/Domain: transmembrane #status predicted <TM7>
 F;211-225/Domain: extracellular #status predicted <EXT4>
 F;226-234/Domain: transmembrane #status predicted <TM8>

   S36348  Length: 234  July 16, 2003 13:42  Type: P  Check: 3344  ..
Found using 'var2' (kam816.key)

...

                                                          |-----|
22      ITHDYPEPTGAKKAQLSTVSDYFRNIRTHSIHPRVSVGYDFGGWRIAADYARYRKWKESN
                                                        72    78

82      SSIKKVTEDIKDNYKETKTEHQENGTFHAVSSLGLSTIYDFQISDKI

...

                        -- Search Statistics --

Times:                      CPU                 Total Elapsed
                         00:00:00.00            00:00:01.00

Number of sequences searched:                   4
Number of sequence hits:                        4
Number of separate matches:                     4
Number of sequence hits saved:                  0
```

```
> O <
O|  |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "var2_pir" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  var2 (AA) ID   var2 AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         wk

Selected files:

        File : var2pir.pep

                    -- Output Parameters --

Format Options:                         File Options:
  Nucleic acid code matching   Exact     Indirect file                 No
  Find non-matching hits only  No        Sequence or key file          No
  Report key used              Yes       List of hits                  Yes
  Note position of hit         Yes       Hit display                   Yes
  Display full annotations     Yes       Name and annotations          Yes
  Sequence context             50

                      -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 b64300 ; TOIG of: b64300  check: 316  from: 1  to: 243
  (from "var2pir.pep")
  TOIG of: b64300  check: 316  from: 1  to: 243

 P1;B64300 - hypothetical protein MJ0002 - Methanococcus jannaschii
 C;Species: Methanococcus jannaschii
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C;Accession: B64300
 R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton,
  G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kerlavage,
  A.R.; Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
  Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Scott, J.L.;
  Geoghagen, N.S.M.; Weidman, J.F.; Fuhrmann, J.L.; Nguyen, D.; Utterback, T.R.;
  Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
  K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.;
  Woese, C.R.; Venter, J.C.
 A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus
  jannaschii.
 A;Reference number: A64300; MUID:96337999; PMID:8688087
 A;Accession: B64300
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-243 <BUL>
 A;Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AAB97989.1;
  PID:g1498759; TIGR:MJ0002
 C;Genetics:
 A;Map position: REV4071-3340

    B64300  Length: 243  July 16, 2003 13:42  Type: P  Check: 316  ..
 Found using 'var2' (kam816.key)

...

                                                                |-----|
62      DFVSFELGKFLRELLKPNANFIEIALSDKVLYSSKYHEDVKGIAYNCICKKLYHHWKGFA
                                                                112   118

122     KPLQKLCEKESYNNPKTLLYILRAYYQGILCLESGEFKSDFSSFRCL

...

-------
1 match found in sequence:
 c83557 ; TOIG of: c83557  check: 491  from: 1  to: 212
  (from "var2pir.pep")
  TOIG of: c83557  check: 491  from: 1  to: 212

 P1;C83557 - chloramphenicol acetyltransferase PA0706 [imported] - Pseudomonas
  aeruginosa (strain PAO1)
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C;Accession: C83557
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey,
  M.J.; Brinkman, F.S.L.; Hufnagle, W.O.; Kowalik, D.J.; Lagrou, M.; Garber,
  R.L.; Goltry, L.; Tolentino, E.; Westbrook-Wadman, S.; Yuan, Y.; Brody, L.L.;
  Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, R.M.; Smith, K.A.;
  Spencer, D.H.; Wong, G.K.S.; Wu, Z.; Paulsen, I.T.; Reizer, J.; Saier, M.H.;
  Hancock, R.E.W.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an
  opportunistic pathogen.
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: C83557
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-212 <STO>
 A;Cross-references: GB:AE004506; GB:AE004091; NID:g9946584; PIDN:AAG04095.1;
  GSPDB:GN00131; PASP:PA0706
 A;Experimental source: strain PAO1
 C;Genetics:
 A;Gene: cat; PA0706
 C;Superfamily: Agrobacterium chloramphenicol acetyltransferase

    C83557  Length: 212  July 16, 2003 13:42  Type: P  Check: 491  ..
 Found using 'var2' (kam816.key)

...

                                                                |-----|
150     YAIVGGNPARTIRKRFSDGDIQNLLEMAWWDWPLADIEAAMPLLCTGDIPALYRHWKQRQ
                                                                200   206

210     ATA

-------
1 match found in sequence:
 i56563 ; TOIG of: i56563  check: 5352  from: 1  to: 896
  (from "var2pir.pep")
  TOIG of: i56563  check: 5352  from: 1  to: 896

 P1;I56563 - interleukin-3 receptor beta-subunit - rat
 C;Species: Rattus sp. (rat)
 C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 23-Jul-1999
 C;Accession: I56563
 R;Appel, K.; Buttini, M.; Sauter, A.; Gebicke-Haerter, P.J.
 J. Neurosci. 15, 5800-5809, 1995
 A;Title: Cloning of rat interleukin-3 receptor beta-subunit from cultured
  microglia and its mRNA expression in vivo.
 A;Reference number: I56563; MUID:95370942; PMID:7643220
 A;Accession: I56563
```

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "varl_pir" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  varl (AA) ID   varl AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         f
2         k or r or h

Selected files:

        File : varlpir.pep

                    -- Output Parameters --

Format Options:                         File Options:
  Nucleic acid code matching   Exact     Indirect file                 No
  Find non-matching hits only  No        Sequence or key file          No
  Report key used              Yes       List of hits                  Yes
  Note position of hit         Yes       Hit display                   Yes
  Display full annotations     Yes       Name and annotations          Yes
  Sequence context             50


                    -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 a71219 ; TOIG of: a71219  check: 5893  from: 1  to: 269
  (from "varlpir.pep")
  TOIG of: a71219  check: 5893  from: 1  to: 269

 P1;A71219 - hypothetical protein PH0012 - Pyrococcus horikoshii
 C;Species: Pyrococcus horikoshii
 C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
 C;Accession: A71219
 R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto,
  S.; Sekine, M.; Baba, S.; Kosugi, H.; Hosoyama, A.; Nagai, Y.; Sakai, M.;
  Ogura, K.; Otsuka, R.; Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.;
  Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.; Aoki, K.;
  Yoshizawa, T.; Nakamura, Y.; Robb, F.T.; Horikoshi, K.; Masuchi, Y.; Shizuya,
  H.; Kikuchi, H.
 DNA Res. 5, 55-76, 1998
 A;Title: Complete sequence and gene organization of the genome of a
  hyper-thermophilic archaebacterium, Pyrococcus horikoshii OT3.
 A;Reference number: A71000; MUID:98344137; PMID:9679194
 A;Accession: A71219
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-269 <KAW>
 A;Cross-references: GB:AP000001; NID:g3236128; PIDN:BAA29080.1; PID:g3256397
 A;Experimental source: strain OT3
 A;Note: this accession replaces an interim accession for a sequence replaced by
  GenBank
 C;Genetics:
 A;Gene: PH0012

    A71219  Length: 269  July 16, 2003 13:37  Type: P  Check: 5893  ..
 Found using 'varl' (kam816.key)

...

                                                                      |-----|
63      FYNKQAWIGLMGVKKAYQRRGIGTEVFRRLLEIGRRKVDTIRLDASSQGYGLYKKFKFVD
                                                                    113   119

123     EYRTVRYELMERPIKRVEGVVEVNKIPNWVKEIDKKAFGDDRIRVLE

...

-------
1 match found in sequence:
 b70452 ; TOIG of: b70452  check: 9383  from: 1  to: 257
  (from "varlpir.pep")
  TOIG of: b70452  check: 9383  from: 1  to: 257

 P1;B70452 - conserved hypothetical protein aq_1768 - Aquifex aeolicus
 C;Species: Aquifex aeolicus
 C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C;Accession: B70452
 R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham,
  D.E.; Overbeek, R.; Snead, M.A.; Keller, M.; Aujay, M.; Huber, R.; Feldman,
  R.A.; Short, J.M.; Olson, G.J.; Swanson, R.V.
 Nature 392, 353-358, 1998
 A;Title: The complete genome of the hyperthermophilic bacterium Aquifex
  aeolicus.
 A;Reference number: A70300; MUID:98196666; PMID:9537320
 A;Accession: B70452
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-257 <AQF>
 A;Cross-references: GB:AE000755; NID:g2984063; PIDN:AAC07614.1; PID:g2984074;
  GB:AE000657
 A;Experimental source: strain VF5
 C;Genetics:
 A;Gene: aq_1768

    B70452  Length: 257  July 16, 2003 13:37  Type: P  Check: 9383  ..
 Found using 'varl' (kam816.key)

...

                                                                      |-----|
94      CIPEENYLKVSSLNPRDTKRLSDKRVERLYLALKPLVKAYCYEIKPEEYNKLYRKFRNLN
                                                                    144   150

154     KMMTHFYKLLIERVKEECGVSEVVVDKYQPSNPFGEDVIFETEAERN

...

-------
1 match found in sequence:
 e71510 ; TOIG of: e71510  check: 2099  from: 1  to: 696
  (from "varlpir.pep")
  TOIG of: e71510  check: 2099  from: 1  to: 696

 P1;E71510 - probable oligopeptide binding lipoprotein - Chlamydia trachomatis
  (serotype D, strain UW3/Cx)
 C;Species: Chlamydia trachomatis
 C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
 C;Accession: E71510
 R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.;
  Mitchell, W.P.; Olinger, L.; Tatusov, R.L.; Zhao, Q.; Koonin, E.V.; Davis, R.W.
 Science 282, 754-759, 1998
 A;Title: Genome sequence of an obligate intracellular pathogen of humans:
  Chlamydia trachomatis.
 A;Reference number: A71570; MUID:99000809; PMID:9784136
 A;Accession: E71510
 A;Status: preliminary
```

```
A;Molecule type: DNA
A;Residues: 1-696 <ARN>
A;Cross-references: GB:AE001321; GB:AE001273; NID:g3328906; PIDN:AAC68080.1;
 PID:g3328915
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: oppA_4

   E71510  Length: 696  July 16, 2003 13:37  Type: P  Check: 2099  ..
Found using 'varl' (kam816.key)

...

                                                                |-----|
576     LGTPPEDPRALWHSEGALEKGSANAVGFCNEEADRIIEQLSYEYDSNKRQALYHRFHEVI
                                                              626   632

636     HEESPYAFLYSRQYSLVYKEFVKNIFVPTEHQDLIPGAQDETVNLSM

...

-------
1 match found in sequence:
 g64502 ; TOIG of: g64502  check: 5094  from: 1  to: 765
  (from "varlpir.pep")
  TOIG of: g64502  check: 5094  from: 1  to: 765

 P1;G64502 - hypothetical protein MJ1626 - Methanococcus jannaschii
 C;Species: Methanococcus jannaschii
 C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Oct-1999
 C;Accession: G64502
 R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton,
  G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton, R.A.; Gocayne, J.D.; Kerlavage,
  A.R.; Dougherty, B.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.;
  Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Scott, J.L.;
  Geoghagen, N.S.M.; Weidman, J.F.; Fuhrmann, J.L.; Nguyen, D.; Utterback, T.R.;
  Kelley, J.M.; Peterson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,
  K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.;
  Woese, C.R.; Venter, J.C.
 A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus
  jannaschii.
 A;Reference number: A64300; MUID:96337999; PMID:8688087
 A;Accession: G64502
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-765 <BUL>
 A;Cross-references: GB:U67602; GB:L77117; NID:g1592214; PIDN:AAB99649.1;
  PID:g1500524; TIGR:MJ1626
 C;Genetics:
 A;Map position: REV1602382-1600085

    G64502  Length: 765  July 16, 2003 13:37  Type: P  Check: 5094  ..
 Found using 'varl' (kam816.key)

...

                                                                |-----|
230     ILANKLYSSEKDERELTINILLNIVDSLTDEQKNILRGHLSVSLLGDEDKKLYRKFKQLF
                                                              280   286

290     EKLDIPAELSDEQIKSLLKSHGKTTLNIILRENIKLPANFYNREFLK

...

-------
1 match found in sequence:
 g97093 ; TOIG of: g97093  check: 8978  from: 1  to: 181
  (from "varlpir.pep")
  TOIG of: g97093  check: 8978  from: 1  to: 181

 P1;G97093 - glutathione peroxidase [imported] - Clostridium acetobutylicum
 C;Species: Clostridium acetobutylicum
 C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
 C;Accession: G97093
 R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson,
  R.; Lee, H.M.; Dubois, J.; Qiu, D.; Hitti, J.; Wolf, Y.I.; Tatusov, R.L.;
  Sabathe, F.; Doucette-Stamm, L.; Soucaille, P.; Daly, M.J.; Bennett, G.N.;
  Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing
  Bacterium Clostridium acetobutylicum.
 A;Reference number: A96900; MUID:21359325; PMID:21359325
 A;Accession: G97093
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-181 <KUR>
 A;Cross-references: GB:AE001437; PIDN:AAK79538.1; PID:g15024524; GSPDB:GN00168
 A;Experimental source: Clostridium acetobutylicum ATCC824
 C;Genetics:
 A;Gene: CAC1571
 C;Superfamily: glutathione peroxidase

    G97093  Length: 181  July 16, 2003 13:37  Type: P  Check: 8978  ..
 Found using 'varl' (kam816.key)

                                                          |-----|
1       MSIYDFKVKDINGEDISMEEYRGKALLIVNTASKCGFTPQYEDLEALYKKFKGENFEVLG
                                                        46    52

61      FPCNQFENQEPGTNNDIKKFCQINYGVTFKIFDKVDVNGENE

...

-------
1 match found in sequence:
 qqbpgl ; TOIG of: qqbpgl  check: 89  from: 1  to: 183
  (from "varlpir.pep")
  TOIG of: qqbpgl  check: 89  from: 1  to: 183

 P1;QQBPGL - git protein - phage lambda
 C;Species: phage lambda
 C;Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 17-Jul-1998
 C;Accession: I43010; C43016; H93737; A04387
 R;Daniels, D.
 submitted to the Nucleic Acid Sequence Database, September 1982
 A;Reference number: A94614
 A;Accession: I43010
 A;Molecule type: DNA
 A;Residues: 1-183 <DAN>
 R;Sanger, F.; Coulson, A.R.; Hong, G.F.; Hill, D.F.; Petersen, G.B.
 J. Mol. Biol. 162, 729-773, 1982
 A;Title: Nucleotide sequence of bacteriophage lambda DNA.
 A;Reference number: A92891; MUID:83189071; PMID:6221115
 A;Accession: C43016
 A;Molecule type: DNA
 A;Residues: 1-183 <SAN>
 A;Cross-references: GB:J02459; GB:M17233; GB:M24325; GB:V00636; GB:X00906;
  NID:g215104
 R;Ineichen, K.; Shepherd, J.C.W.; Bickle, T.A.
 Nucleic Acids Res. 9, 4639-4653, 1981
 A;Title: The DNA sequence of the phage lambda genome between P-L and the gene
  bet.
 A;Reference number: A93737; MUID:82059489; PMID:6458018
 A;Accession: H93737
 A;Molecule type: DNA
 A;Residues: 1-183 <INE>
 A;Cross-references: GB:V00638; NID:g15060
 A;Note: there is no known gene that maps in this region. As this region is not
  overlapped by any other genes and it contains a large open reading frame, the
  authors suggest that it represents a previously unidentified gene, which they
  named git; the propose that, in conjunction with the Ea10 protein, the git gene
  product is involved in production of the Tro phenotype
 A;Note: this phenotype is expressed when phages that possess a mutant cro gene
  and a thermolabile cI repressor gene are unable to propagate at restrictive
```

temperatures; this inability is correlated with the shutoff of host
macromolecular synthesis
A;Note: the authors chose Met-5 as the initial residue because it is preceded
by a Shine-Dalgarno sequence
C;Genetics:
A;Gene: git
A;Map position: 71.10-72.23
C;Superfamily: phage lambda git protein
C;Keywords: transmembrane protein

QQBPGL Length: 183 July 16, 2003 13:37 Type: P Check: 89 ..
Found using 'varl' (kam816.key)

...

```
                                                        |-----|
29      IWVGLLLLSPDNWPEYVNERIGIPHVWHVFVFALAFSLAINVHRLSAIASARYKRFKLRK
                                                       79    85

89      RIKMQNDKVRSVIQNLTEEQSMVLCAALNEGRKYVVTSKQFPYISEL
```

...

---

1 match found in sequence:
t38636 ; TOIG of: t38636 check: 1067 from: 1 to: 1190
(from "varlpir.pep")
TOIG of: t38636 check: 1067 from: 1 to: 1190

P1;T38636 - tat binding homolog - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T38636
R;McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21722
A;Accession: T38636
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1190 <MCL>
A;Cross-references: EMBL:Z98979; PIDN:CAB11703.1; GSPDB:GN00066;
SPDB:SPAC31G5.19
A;Experimental source: strain 972h-; cosmid c31G5
C;Genetics:
A;Gene: SPDB:SPAC31G5.19
A;Map position: 1

T38636 Length: 1190 July 16, 2003 13:37 Type: P Check: 1067 ..
Found using 'varl' (kam816.key)

...

```
                                                        |-----|
762     KRRVLPELPLAPDPPPFTSQKITLKQTKQADMRLLNKLKIKLNALLGSLRARYRKFKKPL
                                                       812   818

822     IDFNDIYCVDPETGHSYRSREECHYEFVDDVVKQIGSDQKFSMMSLE
```

...

-- Search Statistics --

| Times: | CPU | Total Elapsed |
|---|---|---|
| | 00:00:00.00 | 00:00:00.00 |

| | |
|---|---|
| Number of sequences searched: | 7 |
| Number of sequence hits: | 7 |
| Number of separate matches: | 7 |
| Number of sequence hits saved: | 0 |

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "varl_sp" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  varl (AA) ID   varl AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         f
2         k or r or h

Selected files:

        File : varlsp.pep

                    -- Output Parameters --

Format Options:                         File Options:
  Nucleic acid code matching   Exact     Indirect file              No
  Find non-matching hits only  No        Sequence or key file       No
  Report key used              Yes       List of hits               Yes
  Note position of hit         Yes       Hit display                Yes
  Display full annotations     Yes       Name and annotations       Yes
  Sequence context             50

                    -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 calbnaegr ; Calcineurin B subunit (Protein phosphatase 2B regulatory subunit)
  (from "varlsp.pep")
  TOIG of: calb_naegr  check: 936  from: 1  to: 177

 ID   CALB_NAEGR     STANDARD;      PRT;   177 AA.
 AC   P42322;
 DT   01-NOV-1995 (Rel. 32, Created)
 DT   01-NOV-1995 (Rel. 32, Last sequence update)
 DT   16-OCT-2001 (Rel. 40, Last annotation update)
 DE   Calcineurin B subunit (Protein phosphatase 2B regulatory subunit)
 DE   (Calcineurin regulatory subunit).
 GN   CNB1.
 OS   Naegleria gruberi.
 OC   Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
 OX   NCBI_TaxID=5762;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=NEG;
 RX   MEDLINE=95172399; PubMed=7867946;
 RA   Remillard S.P., Lai E.Y., Levy Y.Y., Fulton C.;
 RT   "A calcineurin-B-encoding gene expressed during differentiation of
 RT   the amoeboflagellate Naegleria gruberi contains two introns.";
 RL   Gene 154:39-45(1995).
 CC   -!- FUNCTION: REGULATORY SUBUNIT OF CALCINEURIM, A CALCIUM-DEPENDENT,
 CC       CALMODULIN STIMULATED PROTEIN PHOSPHATASE. CONFERS CALCIUM
 CC       SENSITIVITY (BY SIMILARITY).
 CC   -!- SUBUNIT: COMPOSED OF A CATALYTIC SUBUNIT (A) AND A REGULATORY
 CC       SUBUNIT (B) (BY SIMILARITY).
 CC   -!- MISCELLANEOUS: THIS PROTEIN HAS FOUR FUNCTIONAL CALCIUM-BINDING
 CC       SITES (BY SIMILARITY).
 CC   -!- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; U04380; AAA81896.1; -.
 DR   HSSP; P06705; 1AUI.
 DR   InterPro; IPR002048; EF-hand.
 DR   InterPro; IPR001125; Recoverin.
 DR   Pfam; PF00036; efhand; 4.
 DR   PRINTS; PR00450; RECOVERIN.
 DR   ProDom; PD000012; EF-hand; 2.
 DR   SMART; SM00054; EFh; 4.
 DR   PROSITE; PS00018; EF_HAND; 4.
 KW   Calcium-binding; Repeat.
 FT   CA_BIND      38     49       EF-HAND 1 (BY SIMILARITY).
 FT   CA_BIND      70     81       EF-HAND 2 (BY SIMILARITY).
 FT   CA_BIND     107    118       EF-HAND 3 (BY SIMILARITY).
 FT   CA_BIND     148    159       EF-HAND 4 (BY SIMILARITY).
 SQ   SEQUENCE   177 AA;  20057 MW;  01D4A48E6947E97C CRC64;

  CALB_NAEGR  Length: 177  July 16, 2003 13:39  Type: P  Check: 936  ..
 Found using 'varl' (kam816.key)

                                        |-----|
1       MGTNTSSLRPEEVEEMQKGTNFTQKEIKKLYKRFKKLDKDGNGTISKDEFLMIPELAVNP
                                        29    35

61      LVKRVISIFDENGDGSVNFKEFIAA

...

-------
1 match found in sequence:
 rnh3aquae ; Ribonuclease HIII (EC 3.1.26.-) (RNase HIII).
  (from "varlsp.pep")
  TOIG of: rnh3_aquae  check: 9383  from: 1  to: 257

 ID   RNH3_AQUAE     STANDARD;      PRT;   257 AA.
 AC   O67644;
 DT   30-MAY-2000 (Rel. 39, Created)
 DT   30-MAY-2000 (Rel. 39, Last sequence update)
 DT   15-SEP-2003 (Rel. 42, Last annotation update)
 DE   Ribonuclease HIII (EC 3.1.26.-) (RNase HIII).
 GN   RNHC OR AQ_1768.
 OS   Aquifex aeolicus.
 OC   Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX   NCBI_TaxID=63363;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=VF5;
 RX   MEDLINE=98196666; PubMed=9537320;
 RA   Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA   Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
 RA   Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT   "The complete genome of the hyperthermophilic bacterium Aquifex
 RT   aeolicus.";
 RL   Nature 392:353-358(1998).
 CC   -!- FUNCTION: This enzyme is an endonuclease that degrades the RNA of
 CC       RNA-DNA hybrids specifically (By similarity).
 CC   -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC       phosphomonoester.
 CC   -!- COFACTOR: MAGNESIUM (BY SIMILARITY).
 CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC   -!- SIMILARITY: Belongs to the RNase HII family. RnhC subfamily.
 CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE000755; AAC07614.1; -.
DR   PIR; B70452; B70452.
DR   HAMAP; MF_00053; -; 1.
DR   InterPro; IPR001352; RNase_HII/HIII.
DR   InterPro; IPR004641; RnhIII.
DR   Pfam; PF01351; RNase_HII; 1.
DR   TIGRFAMs; TIGR00716; rnhC; 1.
KW   Hydrolase; Nuclease; Endonuclease; Magnesium; Complete proteome.
FT   ACT_SITE     77     77       BY SIMILARITY.
FT   ACT_SITE    179    179       BY SIMILARITY.
SQ   SEQUENCE   257 AA;  29538 MW;  184FC7B5431F9607 CRC64;

 RNH3_AQUAE  Length: 257  July 16, 2003 13:39  Type: P  Check: 9383  ..
Found using 'var1' (kam816.key)

...

                                                                |-----|
94      CIPEENYLKVSSLNPRDTKRLSDKRVERLYLALKPLVKAYCYEIKPEEYNKLYRKFRNLN
                                                              144   150

154     KMMTHFYKLLIERVKEECGVSEVVVDKYQPSNPFGEDVIFETEAERN

...

-------
1 match found in sequence:
 sieblambd ; Superinfection exclusion protein B.
  (from "var1sp.pep")
  TOIG of: sieb_lambd  check: 89  from: 1  to: 183

 ID   SIEB_LAMBD     STANDARD;      PRT;   183 AA.
 AC   P03762;
 DT   21-JUL-1986 (Rel. 01, Created)
 DT   21-JUL-1986 (Rel. 01, Last sequence update)
 DT   16-OCT-2001 (Rel. 40, Last annotation update)
 DE   Superinfection exclusion protein B.
 GN   SIEB OR GIT.
 OS   Bacteriophage lambda.
 OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 OC   Lambda-like viruses.
 OX   NCBI_TaxID=10710;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RX   MEDLINE=83189071; PubMed=6221115;
 RA   Sanger F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;
 RT   "Nucleotide sequence of bacteriophage lambda DNA.";
 RL   J. Mol. Biol. 162:729-773(1982).
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RX   MEDLINE=82059489; PubMed=6458018;
 RA   Ineichen K., Shepherd J.C.W., Bickle T.A.;
 RT   "The DNA sequence of the phage lambda genome between PL and the gene
 RT   bet.";
 RL   Nucleic Acids Res. 9:4639-4653(1981).
 RN   [3]
 RP   IDENTIFICATION.
 RX   MEDLINE=93328675; PubMed=8335629;
 RA   Ranade K., Poteete A.R.;
 RT   "Superinfection exclusion (sieB) genes of bacteriophages P22 and
 RT   lambda.";
 RL   J. Bacteriol. 175:4712-4718(1993).
 CC   -!- FUNCTION: HAS A ROLE IN THE PREVENTION OF SUPERINFECTION BY
 CC       PHAGES THAT ARE INSENSITIVE TO REPRESSION.
 CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC   -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.
 CC   ---------------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------------
 DR   EMBL; J02459; -; NOT_ANNOTATED_CDS.
 DR   EMBL; V00638; CAA23982.1; ALT_INIT.
 DR   PIR; I43010; QQBPGL.
 KW   Transmembrane.
 FT   TRANSMEM     22     38       POTENTIAL.
 FT   TRANSMEM     54     70       POTENTIAL.
 SQ   SEQUENCE   183 AA;  21426 MW;  F9D1F9CADB886A9A CRC64;

   SIEB_LAMBD  Length: 183  July 16, 2003 13:39  Type: P  Check: 89  ..
 Found using 'var1' (kam816.key)

...

                                                          |-----|
29      IWVGLLLLSPDNWPEYVNERIGIPHVWHVFVFALAFSLAINVHRLSAIASARYKRFKLRK
                                                        79    85

89      RIKMQNDKVRSVIQNLTEEQSMVLCAALNEGRKYVVTSKQFPYISEL

...

-------
1 match found in sequence:
 ye19oceih ; Hypothetical protein OB1419.
  (from "var1sp.pep")
  TOIG of: ye19_oceih  check: 4790  from: 1  to: 88

 ID   YE19_OCEIH     STANDARD;      PRT;    88 AA.
 AC   P59153;
 DT   28-FEB-2003 (Rel. 41, Created)
 DT   28-FEB-2003 (Rel. 41, Last sequence update)
 DT   28-FEB-2003 (Rel. 41, Last annotation update)
 DE   Hypothetical protein OB1419.
 GN   OB1419.
 OS   Oceanobacillus iheyensis.
 OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
 OX   NCBI_TaxID=182710;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=HTE831 / DSM 14371 / JCM 11309;
 RX   MEDLINE=22220767; PubMed=12235376;
 RA   Takami H.,·Takaki Y., Uchiyama I.;
 RT   "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
 RT   Ridge and its unexpected adaptive capabilities to extreme
 RT   environments.";
 RL   Nucleic Acids Res. 30:3927-3935(2002).
 CC   -!- SIMILARITY: BELONGS TO THE UPF0223 FAMILY.
 CC   ---------------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------------
 DR   EMBL; AP004597; BAC13375.1; -.
 DR   HAMAP; MF_01041; -; 1.
 KW   Hypothetical protein; Complete proteome.
 SQ   SEQUENCE   88 AA;  10584 MW;  C27D1A2A0CF3517A CRC64;

  YE19_OCEIH  Length: 88  July 16, 2003 13:39  Type: P  Check: 4790  ..
 Found using 'var1' (kam816.key)
```

```
                                                |-----|
1       MSYQYPMDETWSTEEIIDVVNFFSLIEKAYEKQVDREEILALYRRFKQIVPSKSEEKKLF
                                               41    47

61      TQFQEASGYSSYHVVKKARETNESNIRM

-------
1 match found in sequence:
 yg26metja ; Hypothetical protein MJ1626.
  (from "varlsp.pep")
  TOIG of: yg26_metja  check: 5094  from: 1  to: 765

 ID   YG26_METJA     STANDARD;      PRT;   765 AA.
 AC   Q59019;
 DT   28-FEB-2003 (Rel. 41, Created)
 DT   28-FEB-2003 (Rel. 41, Last sequence update)
 DT   28-FEB-2003 (Rel. 41, Last annotation update)
 DE   Hypothetical protein MJ1626.
 GN   MJ1626.
 OS   Methanococcus jannaschii.
 OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC   Methanocaldococcaceae; Methanocaldococcus.
 OX   NCBI_TaxID=2190;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX   MEDLINE=96337999; PubMed=8688087;
 RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT   jannaschii.";
 RL   Science 273:1058-1073(1996).
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
 CC   the European Bioinformatics Institute.  There are no  restrictions on  its
 CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
 CC   modified and this statement is not removed.  Usage  by  and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; U67602; AAB99649.1; -.
 DR   PIR; G64502; G64502.
 DR   TIGR; MJ1626; -.
 KW   Hypothetical protein; Complete proteome.
 SQ   SEQUENCE   765 AA;  89606 MW;  056EEAA4FEF9B098 CRC64;

  YG26_METJA  Length: 765  July 16, 2003 13:39  Type: P  Check: 5094  ..
 Found using 'var1' (kam816.key)

...

                                                  .   |-----|
230     ILANKLYSSEKDERELTINILLNIVDSLTDEQKNILRGHLSVSLLGDEDKKLYRKFKQLF
                                                      280   286

290     EKLDIPAELSDEQIKSLLKSHGKTTLNIILRENIKLPANFYNREFLK

...


                    -- Search Statistics --

Times:                  CPU                 Total Elapsed
                     00:00:00.00            00:00:01.00
```

```
Number of sequences searched:                5
Number of sequence hits:                     5
Number of separate matches:                  5
Number of sequence hits saved:               0
```

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "var1_spt" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  var1 (AA) ID   var1 AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         f
2         k or r or h

Selected files:

        File : var1spt.pep

                    -- Output Parameters --

Format Options:                          File Options:
  Nucleic acid code matching   Exact      Indirect file                 No
  Find non-matching hits only  No         Sequence or key file          No
  Report key used              Yes        List of hits                  Yes
  Note position of hit         Yes        Hit display                   Yes
  Display full annotations     Yes        Name and annotations          Yes
  Sequence context             50


                    -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 o14114 ; Tat-binding homolog 7, AAA ATPase family protein.
  (from "var1spt.pep")
  TOIG of: o14114  check: 1067  from: 1  to: 1190

 ID   O14114      PRELIMINARY;      PRT;  1190 AA.
 AC   O14114;
 DT   01-JUN-1998 (TrEMBLrel. 06, Created)
 DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Tat-binding homolog 7, AAA ATPase family protein.
 GN   SPAC31G5.19.
 OS   Schizosaccharomyces pombe (Fission yeast).
 OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC   Schizosaccharomyces.
 OX   NCBI_TaxID=4896;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=972h-;
 RA   McLean J., Harris D.;
 RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=972h-;
 RA   Barrell B.G., Rajandream M.A., Wood V.;
 RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; Z98979; CAB11703.1; -.
 DR   GeneDB_SPombe; SPAC31G5.19; -.
 DR   InterPro; IPR003593; AAA_ATPase.
 DR   InterPro; IPR003959; AAA_ATPase_centr.
 DR   InterPro; IPR003960; AAA_sub.
 DR   Pfam; PF00004; AAA; 1.
 DR   SMART; SM00382; AAA; 2.
 DR   PROSITE; PS00674; AAA; 1.
 KW   ATP-binding.
 SQ   SEQUENCE   1190 AA;  135350 MW;  44F5FE817DB188DC CRC64;

   O14114  Length: 1190  July 16, 2003 13:40  Type: P  Check: 1067  ..
 Found using 'var1' (kam816.key)

...

                                                                |-----|
762     KRRVLPELPLAPDPPPFTSQKITLKQTKQADMRLLNKLKIKLNALLGSLRARYRKFKKPL
                                                              812   818

822     IDFNDIYCVDPETGHSYRSREECHYEFVDDVVKQIGSDQKFSMMSLE

...

-------
1 match found in sequence:
 o31001 ; DNA polymerase III holoenzyme tau subunit.
  (from "var1spt.pep")
  TOIG of: o31001  check: 888  from: 1  to: 529

 ID   O31001      PRELIMINARY;      PRT;  529 AA.
 AC   O31001;
 DT   01-JAN-1998 (TrEMBLrel. 05, Created)
 DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   DNA polymerase III holoenzyme tau subunit.
 GN   DNAX.
 OS   Thermus thermophilus.
 OC   Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC   Thermus.
 OX   NCBI_TaxID=274;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=pMF48.KAT;
 RX   MEDLINE=97446162; PubMed=9299346;
 RA   McHenry C.S., Seville M., Cull M.G.;
 RT   "A DNA polymerase III holoenzyme-like subassembly from an extreme
 RT   thermophilic eubacterium.";
 RL   J. Mol. Biol. 272:178-189(1997).
 DR   EMBL; AF025391; AAB82595.1; -.
 DR   HSSP; P28631; 1A5T.
 DR   InterPro; IPR003593; AAA_ATPase.
 DR   InterPro; IPR003959; AAA_ATPase_centr.
 DR   InterPro; IPR001270; Chaprnin_clpA/B.
 DR   InterPro; IPR000862; RFCdomain.
 DR   Pfam; PF00004; AAA; 1.
 DR   PRINTS; PR00300; CLPPROTEASEA.
 DR   SMART; SM00382; AAA; 1.
 KW   ATP-binding.
 SQ   SEQUENCE   529 AA;  58162 MW;  0852FA62DFFEFE33 CRC64;

   O31001  Length: 529  July 16, 2003 13:40  Type: P  Check: 888  ..
 Found using 'var1' (kam816.key)

          |-----|
1       MSALYRRFRPLTFQEVVGQEHVKEPLLKAIREGRLAQAYLFSGPRGVGKTTTARLLAMA
          3     9

...

-------
1 match found in sequence:
 o55649 ; P9-1.
  (from "var1spt.pep")
  TOIG of: o55649  check: 519  from: 1  to: 347
```

```
ID   O55649      PRELIMINARY;      PRT;   347 AA.
AC   O55649;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   P9-1.
OS   Rice black streaked dwarf virus (RBSDV).
OC   Viruses; dsRNA viruses; Reoviridae; Fijivirus.
OX   NCBI_TaxID=10990;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98295849; PubMed=9634092;
RA   Isogai M., Uyeda I., Lee B.-C.;
RT   "Detection and assignment of proteins encoded by rice black streaked
RT   dwarf fijivirus S7, S8, S9, and S10.";
RL   J. Gen. Virol. 79:1487-1494(1998).
DR   EMBL; AB011403; BAA25081.1; -.
SQ   SEQUENCE   347 AA;  39877 MW;  B0DEB74C6A0E7B66 CRC64;

   O55649  Length: 347  July 16, 2003 13:40  Type: P  Check: 519  ..
Found using 'var1' (kam816.key)

...

                                                                |-----|
269      SSLISTPALIREKIAKEGLFKIITSNALRGAPRQTVLFRGINGSESFLNIKRYRKFRTRV
                                                              319   325

329      AGNVDSVIKSDFSSLKLDV

-------
1 match found in sequence:
 o57768 ; Hypothetical protein PH0012.
  (from "varlspt.pep")
  TOIG of: o57768  check: 5893  from: 1  to: 269

 ID   O57768      PRELIMINARY;      PRT;   269 AA.
 AC   O57768;
 DT   01-AUG-1998 (TrEMBLrel. 07, Created)
 DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE   Hypothetical protein PH0012.
 GN   PH0012.
 OS   Pyrococcus horikoshii.
 OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC   Pyrococcus.
 OX   NCBI_TaxID=53953;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=OT3;
 RX   MEDLINE=98344137; PubMed=9679194;
 RA   Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
 RA   Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA   Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA   Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA   Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA   Masuchi Y., Shizuya H., Kikuchi H.;
 RT   "Complete sequence and gene organization of the genome of a hyper-
 RT   thermophilic archaebacterium, Pyrococcus horikoshii OT3.";
 RL   DNA Res. 5:55-76(1998).
 DR   EMBL; AP000001; BAA29080.1; -.
 DR   InterPro; IPR000182; GCN5acetyltransf.
 DR   Pfam; PF00583; Acetyltransf; 1.
 KW   Hypothetical protein; Complete proteome.
 SQ   SEQUENCE   269 AA;  30843 MW;  67B352B36EB6770A CRC64;

    O57768  Length: 269  July 16, 2003 13:40  Type: P  Check: 5893  ..
 Found using 'var1' (kam816.key)

 ...

                                                          |-----|
63       FYNKQAWIGLMGVKKAYQRRGIGTEVFRRLLEIGRRKVDTIRLDASSQGYGLYKKFKFVD
                                                        113   119

123      EYRTVRYELMERPIKRVEGVVEVNKIPNWVKEIDKKAFGDDRIRVLE

...

-------
1 match found in sequence:
 o84486 ; Oligopeptide binding lipoprotein.
  (from "varlspt.pep")
  TOIG of: o84486  check: 2099  from: 1  to: 696

 ID   O84486      PRELIMINARY;      PRT;   696 AA.
 AC   O84486;
 DT   01-NOV-1998 (TrEMBLrel. 08, Created)
 DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   Oligopeptide binding lipoprotein.
 GN   OPPA_4 OR CT480.
 OS   Chlamydia trachomatis.
 OC   Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX   NCBI_TaxID=813;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=D/UW-3/Cx;
 RX   MEDLINE=99000809; PubMed=9784136;
 RA   Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
 RA   Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
 RA   Davis R.W.;
 RT   "Genome sequence of an obligate intracellular pathogen of humans:
 RT   Chlamydia trachomatis.";
 RL   Science 282:754-759(1998).
 DR   EMBL; AE001321; AAC68080.1; -.
 DR   InterPro; IPR000914; SBP_bac_5.
 DR   Pfam; PF00496; SBP_bac_5; 1.
 KW   Lipoprotein; Complete proteome.
 SQ   SEQUENCE   696 AA;  79890 MW;  8531E6CB1D880079 CRC64;

    O84486  Length: 696  July 16, 2003 13:40  Type: P  Check: 2099  ..
 Found using 'var1' (kam816.key)

 ...

                                                          |-----|
576      LGTPPEDPRALWHSEGALEKGSANAVGFCNEEADRIIEQLSYEYDSNKRQALYHRFHEVI
                                                        626   632

636      HEESPYAFLYSRQYSLVYKEFVKNIFVPTEHQDLIPGAQDETVNLSM

...

-------
1 match found in sequence:
 q8bc69 ; Hypothetical protein.
  (from "varlspt.pep")
  TOIG of: q8bc69  check: 1570  from: 1  to: 347

 ID   Q8BC69      PRELIMINARY;      PRT;   347 AA.
 AC   Q8BC69;
 DT   01-MAR-2003 (TrEMBLrel. 23, Created)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Hypothetical protein.
 OS   Rice black streaked dwarf virus (RBSDV).
 OC   Viruses; dsRNA viruses; Reoviridae; Fijivirus.
 OX   NCBI_TaxID=10990;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Henan Zhengzhou;
 RA   Li C., Zhong Y.;
 RT   "Cloning and Expression of s9-1 gene of Rice Black-Streaked Dwarf
 RT   Virus in Escherichia coli.";
```

```
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF540976; AAN28667.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   347 AA;  39854 MW;  65D9703242121D48 CRC64;

   Q8BC69  Length: 347  July 16, 2003 13:40  Type: P  Check: 1570  ..
Found using 'var1' (kam816.key)

...

                                                                |-----|
269     SSLISTPALIREKIAKEGLFKIITSNTLRGASRQTVLFRGINGSESFLNIKRYRKFRTRI
                                                               319   325

329     VGNVDCVIKSDFSSLKLDV

-------
1 match found in sequence:
q8ii55 ; Hypothetical protein.
 (from "varlspt.pep")
 TOIG of: q8ii55  check: 1074  from: 1  to: 1352

ID   Q8II55      PRELIMINARY;      PRT;  1352 AA.
AC   Q8II55;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   PF11_0319.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3D7;
RX   MEDLINE=22255705; PubMed=12368864;
RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA   Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA   Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA   Fraser C.M., Barrell B.;
RT   "Genome sequence of the human malaria parasite Plasmodium
RT   falciparum.";
RL   Nature 419:498-511(2002).
DR   EMBL; AE014840; AAN35903.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   1352 AA;  165114 MW;  85FC8349EFAA8946 CRC64;

  Q8II55  Length: 1352  July 16, 2003 13:40  Type: P  Check: 1074  ..
Found using 'var1' (kam816.key)

...

                                                                |-----|
797     VIMSFYYLSFLKCYNFITYSRLINLILRNVKNEVCIKHLPYVIIGIMNFFKLYKKFKILK
                                                               847   853

857     YKEKEEYKYDENLNKNDLYEQDNISMFNDRYYNKELEKTKINKRKEK

...

-------
1 match found in sequence:
q8izv6 ; Evolutionarily conserved G-patch domain containing protein.
 (from "varlspt.pep")
 TOIG of: q8izv6  check: 1296  from: 1  to: 931

ID   Q8IZV6      PRELIMINARY;      PRT;  931 AA.
AC   Q8IZV6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Evolutionarily conserved G-patch domain containing protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bladder;
RA   Luallen R.J., Sargeant R., Geng Y.;
RT   "Cloning of an Evolutionarily Conserved G-patch Domain Containing
RT   Protein.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF434677; AAN63596.1; -.
SQ   SEQUENCE   931 AA;  103285 MW;  9271E1DFA64906F5 CRC64;

   Q8IZV6  Length: 931  July 16, 2003 13:40  Type: P  Check: 1296  ..
Found using 'var1' (kam816.key)

                                                          |-----|
1       MAARDSDSEEDLVSYGTGLEPLEEGERPKKPIPLQDQTVRDEKGRYKRFHGAFSGGFSAG
                                                         44    50

61      YFNTVGSKEGWTPSTFVSSRQNRADKSVLGPEDFMDEEDL

...

-------
1 match found in sequence:
q8kqx8 ; Sex pilus assembly.
 (from "varlspt.pep")
 TOIG of: q8kqx8  check: 3730  from: 1  to: 95

ID   Q8KQX8      PRELIMINARY;      PRT;    95 AA.
AC   Q8KQX8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Sex pilus assembly.
GN   TRAL.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=666;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22103115; PubMed=12107144;
RA   Beaber J.W., Hochhut B., Waldor M.K.;
RT   "Genomic and Functional Analyses of SXT, an Integrating Antibiotic
RT   Resistance Gene Transfer Element Derived from Vibrio cholerae.";
RL   J. Bacteriol. 184:4259-4269(2002).
DR   EMBL; AY055428; AAL59674.1; -.
SQ   SEQUENCE   95 AA;  10669 MW;  098A0BDD6AC5EC0C CRC64;

   Q8KQX8  Length: 95  July 16, 2003 13:40  Type: P  Check: 3730  ..
Found using 'var1' (kam816.key)

...

                                                            |-----|
4       EPVNIPSYIDDPPHFLLWSADEMAPILLGLVIGIFTGNALVLCLLGLVTTKLYRRFRDGR
                                                           54    60

64      PDGFILHAIYWAGLLPTKAKTIPNPFIRSYLP

-------
1 match found in sequence:
q81x11 ; Ribosomal protein S4.
 (from "varlspt.pep")
 TOIG of: q81x11  check: 7065  from: 1  to: 265
```

```
ID   Q8LX11      PRELIMINARY;      PRT;   265 AA.
AC   Q8LX11;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Ribosomal protein S4.
GN   RPS4.
OS   Laminaria digitata.
OG   Mitochondrion.
OC   Eukaryota; stramenopiles; Phaeophyceae; Laminariales; Laminariaceae;
OC   Laminaria.
OX   NCBI_TaxID=80365;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oudot-le secq M.-P., Kloareg B., Loiseaux-de-Goer S.;
RT   "The mitochondrial genome of the brown alga Laminaria digitata: a
RT   comparative analysis.";
RL   Eur. J. Phycol. 37:163-172(2002).
DR   EMBL; AJ344328; CAC87949.1; -.
DR   InterPro; IPR002942; S4.
DR   Pfam; PF01479; S4; 1.
DR   SMART; SM00363; S4; 1.
KW   Mitochondrion.
SQ   SEQUENCE   265 AA;  31709 MW;  02727BCD1A3E6A7A CRC64;

   Q8LX11  Length: 265  July 16, 2003 13:40  Type: P  Check: 7065  ..
Found using 'varl' (kam816.key)

...

                                                              |-----|
142     HSIYYSRQSIFHNKVLVNGKKIGHGGFILQKGDYVEFCPSQRSAIRARLVARYKRFRSSH
                                                            192   198

202     VKLERWRLSNRFKFELHLHPTPKWVQTDYSNLSFILSSDVCAPVMYP

...

-------
1 match found in sequence:
q8mqt6 ; Protein phosphatase 2B regulatory subunit.
  (from "varlspt.pep")
  TOIG of: q8mqt6  check: 2094  from: 1  to: 177

ID   Q8MQT6      PRELIMINARY;      PRT;   177 AA.
AC   Q8MQT6;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Protein phosphatase 2B regulatory subunit.
GN   CNB.
OS   Toxoplasma gondii.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Sarcocystidae;
OC   Toxoplasma.
OX   NCBI_TaxID=5811;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gaskins E., Kieschnick H., Beckers C.J.;
RT   "Identification of the calcineurin A and B subunits of Toxoplasma
RT   gondii.";
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY125926; AAM97279.1; -.
DR   InterPro; IPR003299; Calflagin.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR001125; Recoverin.
DR   Pfam; PF00036; efhand; 4.
DR   PRINTS; PR01362; CALFLAGIN.
DR   PRINTS; PR00450; RECOVERIN.
DR   ProDom; PD000012; EF-hand; 2.
DR   PROSITE; PS00018; EF_HAND; 4.
SQ   SEQUENCE   177 AA;  19994 MW;  D588EBC014F569E8 CRC64;

   Q8MQT6  Length: 177  July 16, 2003 13:40  Type: P  Check: 2094  ..
Found using 'varl' (kam816.key)

                                        |-----|
1       MGNAQGRLSPQEQMDLLHAANFSERDLKKLYKRFRALDTNQNGELDTHELFDVPELADNP
                                        29    35

61      LVKRVLSIFDTNGDGKVSFVEFLVG

...

-------
1 match found in sequence:
q8n3b7 ; Hypothetical protein (Fragment).
  (from "varlspt.pep")
  TOIG of: q8n3b7  check: 3543  from: 1  to: 932

ID   Q8N3B7      PRELIMINARY;      PRT;   932 AA.
AC   Q8N3B7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP762I1110.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Melanoma;
RA   Blum H., Bauersachs S., Mewes H.W., Weil B., Wiemann S.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL834465; CAD39124.1; -.
DR   InterPro; IPR000467; G_patch.
DR   Pfam; PF01585; G-patch; 1.
DR   PROSITE; PS50174; G_PATCH; 1.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   932 AA;  103518 MW;  C43724C297FD250D CRC64;

   Q8N3B7  Length: 932  July 16, 2003 13:40  Type: P  Check: 3543  ..
Found using 'varl' (kam816.key)

                                                      |-----|
1       DAWARDSDSEEDLVSYGTGLEPLEEGERPKKPIPLQDQTVRDEKGRYKRFHGAFSGGFSA
                                                      45    51

61      GYFNTVGSKEGWTPSTFVSSRQNRADKSVLGPEDFMDEEDL

...

-------
1 match found in sequence:
q8pa45 ; Hypothetical protein XCC1648.
  (from "varlspt.pep")
  TOIG of: q8pa45  check: 6700  from: 1  to: 78

ID   Q8PA45      PRELIMINARY;      PRT;    78 AA.
AC   Q8PA45;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein XCC1648.
GN   XCC1648.
OS   Xanthomonas campestris (pv. campestris).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC   Xanthomonadaceae; Xanthomonas.
OX   NCBI_TaxID=340;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33913 / NCPPB 528;
RX   MEDLINE=22022145; PubMed=12024217;
RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
```

```
RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA   Setubal J.C., Kitajima J.P.;
RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
RT   host specificities.";
RL   Nature 417:459-463(2002).
DR   EMBL; AE012266; AAM40942.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   78 AA;  9080 MW;  BA31765618FD5EFB CRC64;

   Q8PA45  Length: 78  July 16, 2003 13:40  Type: P  Check: 6700  ..
Found using 'varl' (kam816.key)

                    |-----|
1       MPMPSLPPLPSIEPGRYRHFKGGHYEVLDVVRHSETLAPLVLYRPLETDVGMWVRPFEMF
                    15    21

61      VAQVEVEGMMR

...

-------
1 match found in sequence:
 q8rl09 ; TraL.
  (from "varlspt.pep")
  TOIG of: q8rl09  check: 8318  from: 1  to: 93

 ID   Q8RL09      PRELIMINARY;      PRT;    93 AA.
 AC   Q8RL09;
 DT   01-JUN-2002 (TrEMBLrel. 21, Created)
 DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE   TraL.
 OS   Providencia rettgeri.
 OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC   Enterobacteriaceae; Providencia.
 OX   NCBI_TaxID=587;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RA   Boeltner D., MacMahon C., Pembroke J.T., Strike P., Osborn A.M.;
 RT   "R391 is an 89 kb conjugative genomic island comprising elements
 RT   related to plasmids, phages, and transposable elements.";
 RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AY090559; AAM08076.1; -.
 SQ   SEQUENCE   93 AA;  10383 MW;  25DE09E19E787CD9 CRC64;

   Q8RL09  Length: 93  July 16, 2003 13:40  Type: P  Check: 8318  ..
Found using 'varl' (kam816.key)

...

                                                        |-----|
2       EPVSIPSYIDDPPHFLLWSADEMAPILLGLVIGIFTGNALVLCLSGLVTTKLYRRFRDGR
                                                       52    58

62      PDGFILHAIYWAGLLPTKAKTIPNPFIRSYLP

-------
1 match found in sequence:
 q8u487 ; Hypothetical protein PF0204.
  (from "varlspt.pep")
  TOIG of: q8u487  check: 7628  from: 1  to: 369

 ID   Q8U487      PRELIMINARY;      PRT;   369 AA.
 AC   Q8U487;
 DT   01-JUN-2002 (TrEMBLrel. 21, Created)
 DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Hypothetical protein PF0204.
 GN   PF0204.
 OS   Pyrococcus furiosus.
 OC   Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC   Pyrococcus.
 OX   NCBI_TaxID=2261;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA   Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT   "The complete sequence of the Pyrococcus furiosus genome.";
 RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AE010146; AAL80328.1; -.
 DR   InterPro; IPR000583; GATase_2.
 DR   Pfam; PF00310; GATase_2; 1.
 KW   Hypothetical protein; Complete proteome.
 SQ   SEQUENCE   369 AA;  42674 MW;  ACC7975C101F7243 CRC64;

   Q8U487  Length: 369  July 16, 2003 13:40  Type: P  Check: 7628  ..
Found using 'varl' (kam816.key)

...

                                                                |-----|
3       KYRDEKDISACGIFGIMDRKGRRFNGKLVIEAMSQMKVRGNGLGAGFAVYGLYKKFKEYY
                                                               53    59

63      AFHVMIYDKDGREKLEKFLSTNFDIVYVDEIPVNDKAKVIDPPKFMR

...

-------
1 match found in sequence:
 q913e2 ; Hypothetical 40.0 kDa protein.
  (from "varlspt.pep")
  TOIG of: q913e2  check: 1726  from: 1  to: 347

 ID   Q913E2      PRELIMINARY;      PRT;   347 AA.
 AC   Q913E2;
 DT   01-DEC-2001 (TrEMBLrel. 19, Created)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   Hypothetical 40.0 kDa protein.
 OS   Rice black streaked dwarf virus (RBSDV).
 OC   Viruses; dsRNA viruses; Reoviridae; Fijivirus.
 OX   NCBI_TaxID=10990;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=zhjr;
 RA   Bai F., Qu Z., Xu J., Ye M., Shen D.;
 RT   "Rice black streaked dwarf virus segment 9 isolate zhjr.";
 RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AY050487; AAL13113.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   347 AA;  39952 MW;  BF56805FA7489F2B CRC64;

   Q913E2  Length: 347  July 16, 2003 13:40  Type: P  Check: 1726  ..
Found using 'varl' (kam816.key)

...

                                                              |-----|
269     SSLISTPALIREKIAKEGLFKIITSNTLRGASRQTVLFRGINGSESFLNIKRYRKFRTRI
                                                             319   325

329     VGNVDCVIKSDFSSLKLDV
```

```
-------
1 match found in sequence:
 q913e4 ; Hypothetical 39.9 kDa protein.
  (from "var1spt.pep")
  TOIG of: q913e4  check: 1822  from: 1  to: 347

 ID   Q913E4      PRELIMINARY;      PRT;   347 AA.
 AC   Q913E4;
 DT   01-DEC-2001 (TrEMBLrel. 19, Created)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE   Hypothetical 39.9 kDa protein.
 OS   Rice black streaked dwarf virus (RBSDV).
 OC   Viruses; dsRNA viruses; Reoviridae; Fijivirus.
 OX   NCBI_TaxID=10990;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=zhjw;
 RA   Bai F., Qu Z., Xu J., Ye M., Shen D.;
 RT   "Rice black streaked dwarf virus segment 9 isolate zhjw.";
 RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=zhjs;
 RA   Bai F., Qu Z., Yan J., Xu J., Shen D.;
 RT   "Rice black streaked dwarf virus segment 9 isolate zhjs.";
 RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AY050486; AAL13111.1; -.
 DR   EMBL; AF459812; AAL66344.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   347 AA;  39942 MW;  BF5680E237489E34 CRC64;

    Q913E4  Length: 347  July 16, 2003 13:40  Type: P  Check: 1822  ..
 Found using 'var1' (kam816.key)

...

                                                                |-----|
269      SSLISTPALIREKIAKEGLFKIITSNTLRGASRQTVLFRGINGSESFLNIKRYRKFRTRI
                                                              319   325

329      VGNVDCVIKSDFSSLKLDV

-------
1 match found in sequence:
 q91e13 ; Hypothetical protein (40 kDa protein).
  (from "var1spt.pep")
  TOIG of: q91e13  check: 525  from: 1  to: 347

 ID   Q91E13      PRELIMINARY;      PRT;   347 AA.
 AC   Q91E13;
 DT   01-DEC-2001 (TrEMBLrel. 19, Created)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE   Hypothetical protein (40 kDa protein).
 OS   Rice black streaked dwarf virus (RBSDV).
 OC   Viruses; dsRNA viruses; Reoviridae; Fijivirus.
 OX   NCBI_TaxID=10990;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RA   Zhang H., Chen J., Lei J., Adams M.J.;
 RT   "Sequence analysis shows that a dwarfing disease on rice, wheat and
 RT   maize in China is caused by Rice black-streaked dwarf virus (RBSDV).";
 RL   Eur. J. Plant Pathol. 107:563-567(2001).
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RA   Sun L., Fang S., Yu J.;
 RT   "Nucleotide sequence of rice black-streaked dwarf virus fragment 9.";
 RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AJ297430; CAC59968.1; -.
 DR   EMBL; AF536564; AAN04489.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   347 AA;  39962 MW;  2979A7F6226E6917 CRC64;

    Q91E13  Length: 347  July 16, 2003 13:40  Type: P  Check: 525  ..
 Found using 'var1' (kam816.key)

...

                                                                |-----|
269      SSLISTPALIREKIAKEGLFKIITSNTLRGASRQTVLFRGINGSESFLNIKRYRKFRTRI
                                                              319   325

329      VGNVDCVIKSDFSSLKLDV

-------
1 match found in sequence:
 q91e14 ; Hypothetical 40.0 kDa protein.
  (from "var1spt.pep")
  TOIG of: q91e14  check: 9914  from: 1  to: 347

 ID   Q91E14      PRELIMINARY;      PRT;   347 AA.
 AC   Q91E14;
 DT   01-DEC-2001 (TrEMBLrel. 19, Created)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   Hypothetical 40.0 kDa protein.
 OS   Rice black streaked dwarf virus (RBSDV).
 OC   Viruses; dsRNA viruses; Reoviridae; Fijivirus.
 OX   NCBI_TaxID=10990;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RA   Zhang H., Chen J., Lei J., Adams M.J.;
 RT   "Sequence analysis shows that a dwarfing disease on rice, wheat and
 RT   maize in China is caused by Rice black-streaked dwarf virus (RBSDV).";
 RL   Eur. J. Plant Pathol. 107:563-567(2001).
 DR   EMBL; AJ297429; CAC59966.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   347 AA;  39968 MW;  E799BCB488A37B72 CRC64;

    Q91E14  Length: 347  July 16, 2003 13:40  Type: P  Check: 9914  ..
 Found using 'var1' (kam816.key)

...

                                                                |-----|
269      SSLISTPALIREKIAKEGLFKIITSNTLRGASRQTVLFRGINGSESFLNIKRYRKFRTRI
                                                              319   325

329      VGNVDCVIKSDFSSLKLDV

-------
1 match found in sequence:
 q91e20 ; P9-1 protein.
  (from "var1spt.pep")
  TOIG of: q91e20  check: 754  from: 1  to: 347

 ID   Q91E20      PRELIMINARY;      PRT;   347 AA.
 AC   Q91E20;
 DT   01-DEC-2001 (TrEMBLrel. 19, Created)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   P9-1 protein.
 OS   Rice black streaked dwarf virus (RBSDV).
 OC   Viruses; dsRNA viruses; Reoviridae; Fijivirus.
 OX   NCBI_TaxID=10990;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Wuhan S9;
 RA   Zhang H., Chen J., Lei J., Adams M.J.;
 RT   "Sequence analysis shows that a dwarfing disease on rice, wheat and
 RT   maize in China is caused by Rice black-streaked dwarf virus (RBSDV).";
 RL   Eur. J. Plant Pathol. 107:563-567(2001).
 DR   EMBL; AJ291706; CAC59959.1; -.
 SQ   SEQUENCE   347 AA;  39896 MW;  B33A4BF0EED98614 CRC64;
```

```
   Q91E20  Length: 347  July 16, 2003 13:40  Type: P  Check: 754  ..
 Found using 'var1' (kam816.key)

...

                                                                |-----|
269     SSLISTPALIREKIAKEGLFKIITSNTLRGASRQTVLFRGINGSESFLNIKRYRKFRTRI
                                                               319   325

329     VGNVDCVIKSDFSSLKLDV

-------
1 match found in sequence:
 q91h02 ; Hypothetical 39.9 kDa protein.
  (from "var1spt.pep")
  TOIG of: q91h02  check: 183  from: 1  to: 347

 ID   Q91H02      PRELIMINARY;      PRT;   347 AA.
 AC   Q91H02;
 DT   01-DEC-2001 (TrEMBLrel. 19, Created)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   Hypothetical 39.9 kDa protein.
 OS   Rice black streaked dwarf virus (RBSDV).
 OC   Viruses; dsRNA viruses; Reoviridae; Fijivirus.
 OX   NCBI_TaxID=10990;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Shandong;
 RA   Bai F., Qu Z., Xu J., Ye M., Shen D.;
 RT   "Rice black streaked dwarf virus genome segment 9.";
 RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AY039705; AAK73172.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   347 AA;  39897 MW;  6FB9EF7D7C63421B CRC64;

   Q91H02  Length: 347  July 16, 2003 13:40  Type: P  Check: 183  ..
 Found using 'var1' (kam816.key)

...

                                                                |-----|
269     SSLISTPALIREKIAKEGLFKIITSNTLRGASRQTVLFRGINGSESFLNIKRYRKFRTRI
                                                               319   325

329     VGNVDCVIKSDFSSLKLDV

-------
1 match found in sequence:
 q94z19 ; Ribosomal protein S4.
  (from "var1spt.pep")
  TOIG of: q94z19  check: 8109  from: 1  to: 265

 ID   Q94Z19      PRELIMINARY;      PRT;   265 AA.
 AC   Q94Z19;
 DT   01-DEC-2001 (TrEMBLrel. 19, Created)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Ribosomal protein S4.
 GN   RPS4.
 OS   Pylaiella littoralis.
 OG   Mitochondrion.
 OC   Eukaryota; stramenopiles; Phaeophyceae; Ectocarpales;
 OC   Acinetosporaceae; Pylaiella.
 OX   NCBI_TaxID=2885;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RA   Oudot M.P.;
 RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RX   MEDLINE=21372210; PubMed=11479679;
 RA   Oudot M.P., Fontaine J.M., Rousvoal S., Kloareg B.,
 RA   Loiseaux-de Goer S.;
 RT   "The complete sequence of a brown algal mitochondrial genome, the
 RT   Ectocarpale Pylaiella littoralis (L.) Kjellm.";
 RL   J. Mol. Evol. 53:80-88(2001).
 DR   EMBL; AJ277126; CAC50825.1; -.
 DR   InterPro; IPR002942; S4.
 DR   Pfam; PF01479; S4; 1.
 DR   SMART; SM00363; S4; 1.
 KW   Mitochondrion.
 SQ   SEQUENCE   265 AA;  31720 MW;  21EEDF701B6D12F0 CRC64;

   Q94Z19  Length: 265  July 16, 2003 13:40  Type: P  Check: 8109  ..
 Found using 'var1' (kam816.key)

...

                                                                |-----|
142     HSIYYSRQAIQHNKIIVNGKKMGHSGFVLKKGDYVEFCPTQRSAIRARLIARYKRFRSSD
                                                               192   198

202     VRSVRWRLAHRYKLQLQLQPTPKWIQTDYSSLSFLLSSSICPPFMYP

...

-------
1 match found in sequence:
 q97ir9 ; Glutathione peroxidase.
  (from "var1spt.pep")
  TOIG of: q97ir9  check: 8978  from: 1  to: 181

 ID   Q97IR9      PRELIMINARY;      PRT;   181 AA.
 AC   Q97IR9;
 DT   01-OCT-2001 (TrEMBLrel. 18, Created)
 DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE   Glutathione peroxidase.
 GN   CAC1571.
 OS   Clostridium acetobutylicum.
 OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC   Clostridium.
 OX   NCBI_TaxID=1488;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX   MEDLINE=21359325; PubMed=11466286;
 RA   Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA   Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA   Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA   Bennett G.N., Koonin E.V., Smith D.R.;
 RT   "Genome sequence and comparative analysis of the solvent-producing
 RT   bacterium Clostridium acetobutylicum.";
 RL   J. Bacteriol. 183:4823-4838(2001).
 DR   EMBL; AE007667; AAK79538.1; -.
 DR   InterPro; IPR000889; Glut_peroxidase.
 DR   Pfam; PF00255; GSHPx; 1.
 DR   PRINTS; PR01011; GLUTPROXDASE.
 DR   PROSITE; PS00460; GLUTATHIONE_PEROXID_1; 1.
 DR   PROSITE; PS00763; GLUTATHIONE_PEROXID_2; 1.
 KW   Peroxidase; Complete proteome.
 SQ   SEQUENCE   181 AA;  20870 MW;  D688BDDC0C27D49C CRC64;

   Q97IR9  Length: 181  July 16, 2003 13:40  Type: P  Check: 8978  ..
 Found using 'var1' (kam816.key)

                                                          |-----|
1       MSIYDFKVKDINGEDISMEEYRGKALLIVNTASKCGFTPQYEDLEALYKKFKGENFEVLG
                                                         46    52

61      FPCNQFENQEPGTNNDIKKFCQINYGVTFKIFDKVDVNGENE

...

-------
```

```
1 match found in sequence:
 q9brr8 ; Similar to RIKEN cDNA 1300003A17 gene.
  (from "varlspt.pep")
  TOIG of: q9brr8  check: 1310  from: 1  to: 931

 ID   Q9BRR8      PRELIMINARY;      PRT;   931 AA.
 AC   Q9BRR8;
 DT   01-JUN-2001 (TrEMBLrel. 17, Created)
 DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Similar to RIKEN cDNA 1300003A17 gene.
 OS   Homo sapiens (Human).
 OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX   NCBI_TaxID=9606;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   TISSUE=Muscle;
 RA   Strausberg R.;
 RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; BC006108; AAH06108.1; -.
 DR   InterPro; IPR000467; G_patch.
 DR   Pfam; PF01585; G-patch; 1.
 DR   SMART; SM00443; G_patch; 1.
 DR   PROSITE; PS50174; G_PATCH; 1.
 SQ   SEQUENCE   931 AA;  103345 MW;  407082BEC44904F4 CRC64;

    Q9BRR8  Length: 931  July 16, 2003 13:40  Type: P  Check: 1310  ..
 Found using 'varl' (kam816.key)

                                                    |-----|
1       MAARDSDSEEDLVSYGTGLEPLEEGERPKKPIPLQDQTVRDEKGRYKRFHGAFSGGFSAG
                                                   44    50

61      YFNTVGSKEGWTPSTFVSSRQNRADKSVLGPEDFMDEEDL

...

-------
1 match found in sequence:
 q9dbm1 ; 1300003A17Rik protein.
  (from "varlspt.pep")
  TOIG of: q9dbm1  check: 1796  from: 1  to: 930

 ID   Q9DBM1      PRELIMINARY;      PRT;   930 AA.
 AC   Q9DBM1;
 DT   01-JUN-2001 (TrEMBLrel. 17, Created)
 DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   1300003A17Rik protein.
 GN   1300003A17RIK.
 OS   Mus musculus (Mouse).
 OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX   NCBI_TaxID=10090;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=C57BL/6J; TISSUE=Liver;
 RX   MEDLINE=21085660; PubMed=11217851;
 RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA   Hayashizaki Y.;
 RT   "Functional annotation of a full-length mouse cDNA collection.";
 RL   Nature 409:685-690(2001).
 DR   EMBL; AK004868; BAB23629.1; -.
 DR   MGD; MGI:1914721; 1300003A17Rik.
 DR   InterPro; IPR000467; G_patch.
 DR   Pfam; PF01585; G-patch; 1.
 DR   SMART; SM00443; G_patch; 1.
 DR   PROSITE; PS50174; G_PATCH; 1.
 SQ   SEQUENCE   930 AA;  103007 MW;  62D383B72424EF73 CRC64;

    Q9DBM1  Length: 930  July 16, 2003 13:40  Type: P  Check: 1796  ..
 Found using 'varl' (kam816.key)

                                                    |-----|
1       MAALDSDSDEDLISYGTGLEPLDEGERPKKPIPLQDQTVRDEKGRYKRFHGAFSGGFSAG
                                                   44    50

61      YFNTVGSKEGWTPSTFVSSRQNRADKSALGPEDFMDEEDL

...

-------
1 match found in sequence:
 q9p7b3 ; Calcium permease family membrane transporter.
  (from "varlspt.pep")
  TOIG of: q9p7b3  check: 6779  from: 1  to: 881

 ID   Q9P7B3      PRELIMINARY;      PRT;   881 AA.
 AC   Q9P7B3;
 DT   01-OCT-2000 (TrEMBLrel. 15, Created)
 DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Calcium permease family membrane transporter.
 GN   SPAC521.04C.
 OS   Schizosaccharomyces pombe (Fission yeast).
 OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC   Schizosaccharomyces.
 OX   NCBI_TaxID=4896;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=972h-;
 RA   Brown S., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
 RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AL163481; CAB86468.1; -.
 DR   GeneDB_SPombe; SPAC521.04c; -.
 DR   InterPro; IPR005185; DUF307.
 DR   InterPro; IPR001092; HLH_basic.
 DR   InterPro; IPR004837; NaCa_Exmemb.
 DR   Pfam; PF03733; DUF307; 1.
 DR   Pfam; PF01699; Na_Ca_Ex; 2.
 DR   PROSITE; PS00038; HLH_1; 1.
 SQ   SEQUENCE   881 AA;  98857 MW;  4190EBFD883196C4 CRC64;

    Q9P7B3  Length: 881  July 16, 2003 13:40  Type: P  Check: 6779  ..
 Found using 'varl' (kam816.key)

...

                                                                |-----|
118     IPKQNFKVRPLSPLRGQSPASLRSEESFTLKERQNAINKTRAFGMRLWKPALYKKFRSIN
                                                               168   174

178     RDADIDIHDEPLKRPNTSISNVIWLICFGAPLFLVIFICYIFFTVLS

...

-------
1 match found in sequence:
```

```
q9rp72 ; Hypothetical 81.0 kDa protein.
 (from "varlspt.pep")
 TOIG of: q9rp72  check: 2853  from: 1  to: 706

ID   Q9RP72      PRELIMINARY;      PRT;   706 AA.
AC   Q9RP72;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Hypothetical 81.0 kDa protein.
OS   Pasteurella haemolytica.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=75985;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21295097; PubMed=11401986;
RA   Lo R.Y.C., McKerral L.J., Hills T.L., Kostrzynska M.;
RT   "Analysis of the Capsule Biosynthetic Locus of Mannheimia
RT   (Pasteurella) haemolytica A1 and Proposal of a Nomenclature System.";
RL   Infect. Immun. 69:4458-4464(2001).
DR   EMBL; AF170495; AAF08247.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   706 AA;  81021 MW;  D80B20C2145F0D37 CRC64;

   Q9RP72  Length: 706  July 16, 2003 13:40  Type: P  Check: 2853  ..
Found using 'varl' (kam816.key)

...

                                                              |-----|
342     KKQMDMLLPALLAYNGCIYDRASKDTSGKYAYPEQYTHLIREGVNFNEIVKLYKKFKVFL
                                                            392   398

402     NVNTITQSTTMMSRRVYELLASGTPVVSTPSKAITAQFPGIVITVNN

...


                     -- Search Statistics --

Times:                    CPU                 Total Elapsed
                      00:00:00.00             00:00:00.00

Number of sequences searched:                    27
Number of sequence hits:                         27
Number of separate matches:                      27
Number of sequence hits saved:                    0
```

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "var2_sp" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  var2 (AA) ID   var2 AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         wk

Selected files:

       File : var2sp.pep

                    -- Output Parameters --

Format Options:                        File Options:
  Nucleic acid code matching   Exact     Indirect file                No
  Find non-matching hits only  No        Sequence or key file         No
  Report key used              Yes       List of hits                 Yes
  Note position of hit         Yes       Hit display                  Yes
  Display full annotations     Yes       Name and annotations         Yes
  Sequence context             50


                      -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 cat4pseae ; Chloramphenicol acetyltransferase (EC 2.3.1.28) (Xenobiotic
  (from "var2sp.pep")
  TOIG of: cat4_pseae  check: 491  from: 1  to: 212

 ID   CAT4_PSEAE     STANDARD;      PRT;   212 AA.
 AC   P26841; Q9S542;
 DT   01-AUG-1992 (Rel. 23, Created)
 DT   16-OCT-2001 (Rel. 40, Last sequence update)
 DT   28-FEB-2003 (Rel. 41, Last annotation update)
 DE   Chloramphenicol acetyltransferase (EC 2.3.1.28) (Xenobiotic
 DE   acetyltransferase) (XAT).
 GN   CAT OR CATB7 OR PA0706.
 OS   Pseudomonas aeruginosa.
 OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC   Pseudomonadaceae; Pseudomonas.
 OX   NCBI_TaxID=287;
 RN   [1]
 RP   SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC   STRAIN=PAO222;
 RX   MEDLINE=99290057; PubMed=10361706;
 RA   White P.A., Stokes H.W., Bunny K.L., Hall R.M.;
 RT   "Characterisation of a chloramphenicol acetyltransferase determinant
 RT   found in the chromosome of Pseudomonas aeruginosa.";
 RL   FEMS Microbiol. Lett. 175:27-35(1999).
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=ATCC 15692 / PAO1;
 RX   MEDLINE=20437337; PubMed=10984043;
 RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
 RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 RT   opportunistic pathogen.";
 RL   Nature 406:959-964(2000).
 RN   [3]
 RP   SEQUENCE FROM N.A., AND CRYSTALLIZATION.
 RC   STRAIN=ATCC 29260 / PA103;
 RX   MEDLINE=97332513; PubMed=9188747;
 RA   Tian Y., Beaman T.W., Roderick S.L.;
 RT   "Purification and crystallization of Pseudomonas aeruginosa
 RT   chloramphenicol acetyltransferase.";
 RL   Proteins 28:298-300(1997).
 RN   [4]
 RP   SEQUENCE OF 1-64 FROM N.A.
 RC   STRAIN=ATCC 29260 / PA103;
 RX   MEDLINE=88262583; PubMed=3133641;
 RA   Hindahl M.S., Frank D.W., Hamood A., Iglewski B.H.;
 RT   "Characterization of a gene that regulates toxin A synthesis in
 RT   Pseudomonas aeruginosa.";
 RL   Nucleic Acids Res. 16:5699-5699(1988).
 RN   [5]
 RP   X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).
 RC   STRAIN=ATCC 29260 / PA103;
 RX   MEDLINE=98244846; PubMed=9578552;
 RA   Beaman T.W., Sugantino M., Roderick S.L.;
 RT   "Structure of the hexapeptide xenobiotic acetyltransferase from
 RT   Pseudomonas aeruginosa.";
 RL   Biochemistry 37:6689-6696(1998).
 CC   -!- FUNCTION: THIS ENZYME IS AN EFFECTOR OF CHLORAMPHENICOL (CM)
 CC       RESISTANCE IN BACTERIA. ACETYLATES CM BUT NOT 1-ACETOXY-CM.
 CC   -!- CATALYTIC ACTIVITY: Acetyl-CoA + chloramphenicol = CoA +
 CC       chloramphenicol 3-acetate.
 CC   -!- SUBUNIT: Homotrimer.
 CC   -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NODL FAMILY OF
 CC       ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).
 CC   -!- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN IN POSITIONS 11
 CC       TO 37 DUE TO A FRAMESHIFT.
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; AF036933; AAD02068.1; -.
 DR   EMBL; AE004506; AAG04095.1; -.
 DR   EMBL; X12366; -; NOT_ANNOTATED_CDS.
 DR   PIR; C83557; C83557.
 DR   PDB; 1XAT; 17-JUN-98.
 DR   PDB; 2XAT; 17-JUN-98.
 DR   InterPro; IPR001451; Hexapep_transf.
 DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; FALSE_NEG.
 KW   Antibiotic resistance; Transferase; Acyltransferase; Repeat;
 KW   3D-structure; Complete proteome.
 FT   ACT_SITE     79     79
 FT   CONFLICT     97     97       V -> A (IN REF. 3).
 FT   CONFLICT    116    116       D -> E (IN REF. 3).
 FT   CONFLICT    203    203       R -> Q (IN REF. 3).
 FT   TURN          8      9
 FT   STRAND       12     13
 FT   TURN         15     17
 FT   TURN         21     22
 FT   STRAND       23     25
 FT   TURN         27     28
 FT   STRAND       30     32
 FT   HELIX        34     36
 FT   HELIX        40     43
```

```
FT   STRAND       44     45
FT   TURN         46     46
FT   STRAND       57     59
FT   STRAND       64     66
FT   TURN         67     68
FT   STRAND       70     72
FT   TURN         75     78
FT   TURN         81     82
FT   HELIX        89     91
FT   TURN         93     94
FT   HELIX        96     98
FT   STRAND      111    113
FT   STRAND      118    119
FT   TURN        121    122
FT   STRAND      124    125
FT   TURN        127    128
FT   STRAND      130    131
FT   TURN        133    134
FT   STRAND      136    137
FT   TURN        139    140
FT   STRAND      142    143
FT   STRAND      147    147
FT   TURN        149    150
FT   STRAND      151    154
FT   TURN        155    158
FT   STRAND      159    163
FT   HELIX       167    176
FT   HELIX       178    180
FT   HELIX       183    193
FT   TURN        194    194
FT   HELIX       198    208
FT   TURN        209    209
SQ   SEQUENCE   212 AA;  23524 MW;  306EEF39978C19FC CRC64;

 CAT4_PSEAE  Length: 212  July 16, 2003 13:42  Type: P  Check: 491  ..
Found using 'var2' (kam816.key)

...

                                                                |-----|
150     YAIVGGNPARTIRKRFSDGDIQNLLEMAWWDWPLADIEAAMPLLCTGDIPALYRHWKQRQ
                                                              200   206

210     ATA

-------
1 match found in sequence:
 op65neigo ; Opacity protein OPA65 precursor (Fragment).
  (from "var2sp.pep")
  TOIG of: op65_neigo  check: 3344  from: 1  to: 234

 ID   OP65_NEIGO     STANDARD;      PRT;   234 AA.
 AC   Q04885;
 DT   01-OCT-1993 (Rel. 27, Created)
 DT   01-OCT-1993 (Rel. 27, Last sequence update)
 DT   01-NOV-1995 (Rel. 32, Last annotation update)
 DE   Opacity protein OPA65 precursor (Fragment).
 OS   Neisseria gonorrhoeae.
 OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC   Neisseriaceae; Neisseria.
 OX   NCBI_TaxID=485;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=VP1;
 RX   MEDLINE=93178439; PubMed=8440254;
 RA   Kupsch E.-M., Knepper B., Kuroki T., Heuer I., Meyer T.F.;
 RT   "Variable opacity (Opa) outer membrane proteins account for the cell
 RT   tropisms displayed by Neisseria gonorrhoeae for human leukocytes and
 RT   epithelial cells.";
 RL   EMBO J. 12:641-650(1993).
 CC   -!- FUNCTION: IMPLICATED IN A NUMBER OF ADHERENCE FUNCTIONS. OPA
 CC       PROTEINS ARE IMPLICATED IN PATHOGENESIS AND ARE SUBJECT TO PHASE
 CC       VARIATION.
 CC   -!- SUBCELLULAR LOCATION: Outer membrane.
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; Z18940; CAA79373.1; -.
 DR   PIR; S36348; S36348.
 DR   InterPro; IPR003394; Porin_opacity.
 DR   Pfam; PF02462; Opacity; 1.
 KW   Outer membrane; Multigene family; Signal.
 FT   NON_TER       1      1
 FT   SIGNAL       <1      1       POTENTIAL.
 FT   CHAIN         2    >234      OPACITY PROTEIN OPA65.
 FT   NON_TER     234    234
 SQ   SEQUENCE   234 AA;  26242 MW;  D66A0BAA8424C2F1 CRC64;

  OP65_NEIGO  Length: 234  July 16, 2003 13:42  Type: P  Check: 3344  ..
 Found using 'var2' (kam816.key)

...

                                                                |-----|
22      ITHDYPEPTGAKKAQLSTVSDYFRNIRTHSIHPRVSVGYDFGGWRIAADYARYRKWKESN
                                                              72    78

82      SSIKKVTEDIKDNYKETKTEHQENGTFHAVSSLGLSTIYDFQISDKI

...

-------
1 match found in sequence:
 y002metja ; Hypothetical protein MJ0002.
  (from "var2sp.pep")
  TOIG of: y002_metja  check: 316  from: 1  to: 243

 ID   Y002_METJA     STANDARD;      PRT;   243 AA.
 AC   Q60312;
 DT   01-NOV-1997 (Rel. 35, Created)
 DT   01-NOV-1997 (Rel. 35, Last sequence update)
 DT   16-OCT-2001 (Rel. 40, Last annotation update)
 DE   Hypothetical protein MJ0002.
 GN   MJ0002.
 OS   Methanococcus jannaschii.
 OC   Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC   Methanocaldococcaceae; Methanocaldococcus.
 OX   NCBI_TaxID=2190;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX   MEDLINE=96337999; PubMed=8688087;
 RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA   Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT   jannaschii.";
 RL   Science 273:1058-1073(1996).
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; U67459; AAB97989.1; -.
DR   PIR; B64300; B64300.
DR   TIGR; MJ0002; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   243 AA;  28492 MW;  2B1BE2C59F816CB9 CRC64;

 Y002_METJA  Length: 243  July 16, 2003 13:42  Type: P  Check: 316  ..
Found using 'var2' (kam816.key)

...

                                                              |-----|
62      DFVSFELGKFLRELLKPNANFIEIALSDKVLYSSKYHEDVKGIAYNCICKKLYHHWKGFA
                                                             112   118

122     KPLQKLCEKESYNNPKTLLYILRAYYQGILCLESGEFKSDFSSFRCL

...


                     -- Search Statistics --

Times:                   CPU                  Total Elapsed
                    00:00:00.00               00:00:00.00

Number of sequences searched:                      3
Number of sequence hits:                           3
Number of separate matches:                        3
Number of sequence hits saved:                     0
```

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "var2_spt" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  var2 (AA) ID   var2 AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         wk

Selected files:

        File : var2spt.pep

                    -- Output Parameters --

Format Options:                         File Options:
  Nucleic acid code matching   Exact     Indirect file              No
  Find non-matching hits only  No        Sequence or key file       No
  Report key used              Yes       List of hits               Yes
  Note position of hit         Yes       Hit display                Yes
  Display full annotations     Yes       Name and annotations       Yes
  Sequence context             50


                     -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 o97035 ; PLC-betaH1.
  (from "var2spt.pep")
  TOIG of: o97035  check: 1081  from: 1  to: 1012

 ID   O97035      PRELIMINARY;      PRT;  1012 AA.
 AC   O97035;
 DT   01-MAY-1999 (TrEMBLrel. 10, Created)
 DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   PLC-betaH1.
 OS   Hydra magnipapillata (Hydra).
 OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
 OC   Hydridae; Hydra.
 OX   NCBI_TaxID=6085;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=105;
 RA   Koyanagi M., Ono K., Suga H., Iwabe N., Miyata T.;
 RT   "Phospholipase C cDNAs from sponge and hydra: Antiquity of genes
 RT   involved in the inositol phospholipid signaling pathway.";
 RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
 CC   -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 DR   EMBL; AB017511; BAA76276.1; -.
 DR   HSSP; P10688; 1DJX.
 DR   InterPro; IPR000008; C2.
 DR   InterPro; IPR001192; PI_PLC.
 DR   InterPro; IPR000909; PI_PLC_Xdom.
 DR   InterPro; IPR001711; PI_PLC_Y.
 DR   Pfam; PF00168; C2; 1.
 DR   Pfam; PF00388; PI-PLC-X; 1.
 DR   Pfam; PF00387; PI-PLC-Y; 1.
 DR   PRINTS; PR00390; PHPHLIPASEC.
 DR   ProDom; PD001202; PI_PLC_Y; 1.
 DR   SMART; SM00239; C2; 1.
 DR   SMART; SM00148; PLCXc; 1.
 DR   SMART; SM00149; PLCYc; 1.
 DR   PROSITE; PS50004; C2_DOMAIN_2; 1.
 DR   PROSITE; PS50007; PIPLC_X_DOMAIN; 1.
 DR   PROSITE; PS50008; PIPLC_Y_DOMAIN; 1.
 SQ   SEQUENCE   1012 AA;  116186 MW;  BD7CF51F96A23118 CRC64;

   O97035  Length: 1012  July 16, 2003 13:43  Type: P  Check: 1081  ..
 Found using 'var2' (kam816.key)

...

                                                                |-----|
886     KRSQSQSLGRLESPKQPIRRTFSTPKLLGEDTSSDYLKISGLSINTLKEHKLYKKWKLKF
                                                              936   942

946     DSDFLNLMQKQAKGRIQCRKEIEAKTKKLTLNLEKKNQLYKRIKTNF

...

-------
1 match found in sequence:
 q64146 ; Interleukin-3 receptor beta-subunit (Fragment).
  (from "var2spt.pep")
  TOIG of: q64146  check: 5352  from: 1  to: 896

 ID   Q64146      PRELIMINARY;      PRT;   896 AA.
 AC   Q64146;
 DT   01-NOV-1996 (TrEMBLrel. 01, Created)
 DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Interleukin-3 receptor beta-subunit (Fragment).
 GN   RIL-3R<BETA>.
 OS   Rattus norvegicus (Rat).
 OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX   NCBI_TaxID=10116;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RX   MEDLINE=95370942; PubMed=7643220;
 RA   Appel K., Buttini M., Sauter A., Gebicke-Haerter P.J.;
 RT   "Cloning of rat interleukin-3 receptor beta-subunit from cultured
 RT   microglia and its mRNA expression in vivo.";
 RL   J. Neurosci. 15:5800-5809(1995).
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=WISTAR; TISSUE=Brain;
 RA   Gebicke-Haerter P.J.;
 RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; S79263; AAB35068.1; -.
 DR   EMBL; AJ000555; CAA04186.1; -.
 DR   InterPro; IPR002996; CR1A.
 DR   InterPro; IPR000282; Cytok_receptor_2.
 DR   InterPro; IPR003961; FN_III.
 DR   InterPro; IPR003531; Hemtopoptn_S_F1.
 DR   Pfam; PF00041; fn3; 2.
 DR   SMART; SM00060; FN3; 2.
 DR   PROSITE; PS01355; HEMATOPO_REC_S_F1; 1.
 KW   Signal.
 FT   NON_TER       1      1
 FT   NON_TER     896    896
 SQ   SEQUENCE   896 AA;  99504 MW;  4C6E3B288A4A1052 CRC64;

    Q64146  Length: 896  July 16, 2003 13:43  Type: P  Check: 5352  ..
 Found using 'var2' (kam816.key)

...
```

```
                                                                |-----|
421     PEYKGLWSEWSNECTWTTDWVMPTLWIVLILVFLILTFLLALRFGCIYGCKLYRRWKEKI
                                                              471   477

481     PNPSKSLLFQDGGKGLWPPGSTVTFSSKNPTPQGPQNNLFSELQGVS

...

-------
1 match found in sequence:
 q8jtu9 ; Hypothetical protein.
  (from "var2spt.pep")
  TOIG of: q8jtu9  check: 6954  from: 1  to: 437

 ID   Q8JTU9      PRELIMINARY;      PRT;   437 AA.
 AC   Q8JTU9;
 DT   01-OCT-2002 (TrEMBLrel. 22, Created)
 DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE   Hypothetical protein.
 GN   LW054.
 OS   Lumpy skin disease virus (LSDV).
 OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC   Capripoxvirus.
 OX   NCBI_TaxID=59509;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Neethling vaccine LW 1959;
 RA   Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Stipinovich C.,
 RA   Lu Z., Vreede F.T., Taljaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;
 RT   "Molecular characterization of the South African vaccine strain and
 RT   the field isolate of lumpy skin disease virus.";
 RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AF409138; AAN02779.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   437 AA;  50521 MW;  500D63C6D9CC08EC CRC64;

    Q8JTU9  Length: 437  July 16, 2003 13:43  Type: P  Check: 6954  ..
 Found using 'var2' (kam816.key)

...

                                                                |-----|
274     IIRNYSRKETDRYINAEKIIHFINKYSCLNESIYNEVPPESITVQEFIFSALYHKWKKFD
                                                              324   330

334     DSLLKGISLCCSLICVLKPKKDIKRNEISKLCNIINNDSNKKSSLNN

...

-------
1 match found in sequence:
 q8v3p3 ; SPV051 hypothetical protein.
  (from "var2spt.pep")
  TOIG of: q8v3p3  check: 8140  from: 1  to: 439

 ID   Q8V3P3      PRELIMINARY;      PRT;   439 AA.
 AC   Q8V3P3;
 DT   01-MAR-2002 (TrEMBLrel. 20, Created)
 DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE   SPV051 hypothetical protein.
 GN   SPV051.
 OS   Swinepox virus.
 OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC   Suipoxvirus.
 OX   NCBI_TaxID=10276;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=17077-99;
 RX   MEDLINE=21624277; PubMed=11752168;
 RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Osorio F.A., Balinsky C.,
 RA   Kutish G.F., Rock D.L.;
 RT   "The genome of swinepox virus.";
 RL   J. Virol. 76:783-790(2002).
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=17077-99;
 RA   Afonso C.L., Tulman E.R., Lu Z., Balinsky C., Osorio F.A., Zsak L.,
 RA   Kutish G.F., Rock D.L.;
 RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AF410153; AAL69790.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   439 AA;  50984 MW;  4991293795BF3850 CRC64;

    Q8V3P3  Length: 439  July 16, 2003 13:43  Type: P  Check: 8140  ..
 Found using 'var2' (kam816.key)

...

                                                                |-----|
275     VRNYTKKSTDKYINVQEIIDFIDKYSSSDDSVYSNTLPEKCVTIQEFMFSALYKKWKTFD
                                                              325   331

335     DTLLRGVSICSSLICILSPKKAISNDDIDILCKIINKNNKNKNDIIN

...

-------
1 match found in sequence:
 q91mv3 ; LSDV054 hypothetical protein.
  (from "var2spt.pep")
  TOIG of: q91mv3  check: 7600  from: 1  to: 437

 ID   Q91MV3      PRELIMINARY;      PRT;   437 AA.
 AC   Q91MV3;
 DT   01-DEC-2001 (TrEMBLrel. 19, Created)
 DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE   LSDV054 hypothetical protein.
 GN   LSDV054 OR LD054.
 OS   Lumpy skin disease virus (LSDV).
 OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC   Capripoxvirus.
 OX   NCBI_TaxID=59509;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Neethling 2490;
 RX   MEDLINE=21329495; PubMed=11435593;
 RA   Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
 RT   "Genome of lumpy skin disease virus.";
 RL   J. Virol. 75:7122-7130(2001).
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Neethling 2490;
 RA   Tulman E.R., Afonso C.L., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
 RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
 RN   [3]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Neethling Warmbaths LW;
 RA   Kara P.D., Afonso C.L., Wallace D.B., Kutish G.F., Stipinovich C.,
 RA   Lu Z., Vreede F.T., Taljaard L.C.F., Zsak A., Viljoen G.J., Rock D.L.;
 RT   "Molecular characterization of the South African vaccine strain and
 RT   the field isolate of lumpy skin disease virus.";
 RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AF325528; AAK85015.1; -.
 DR   EMBL; AF409137; AAN02622.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   437 AA;  50555 MW;  2F9798DBC80C223F CRC64;

    Q91MV3  Length: 437  July 16, 2003 13:43  Type: P  Check: 7600  ..
 Found using 'var2' (kam816.key)

...

                                                                |-----|
```

```
274     IIRNYSRKDTDRYINAEKIIHFINKYSCLNESIYNEVPPESITVQEFIFSALYHKWKKFD
                                                          324   330

334     DSLLKGISLCCSLICVLKPKKDIKRNEISKLCNIINNDSNKKSSLNN

...

-------
1 match found in sequence:
 q9dfi1 ; Doublesex-like protein Dmrt11 (Fragment).
  (from "var2spt.pep")
  TOIG of: q9dfi1  check: 8097  from: 1  to: 31

 ID   Q9DFI1      PRELIMINARY;      PRT;    31 AA.
 AC   Q9DFI1;
 DT   01-MAR-2001 (TrEMBLrel. 16, Created)
 DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Doublesex-like protein Dmrt11 (Fragment).
 OS   Brachydanio rerio (Zebrafish) (Danio rerio).
 OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC   Cyprinidae; Danio.
 OX   NCBI_TaxID=7955;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RA   Huang X., Zhou R.;
 RT   "A conserved family of doublesex related genes from fish.";
 RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AF272963; AAG18559.1; -.
 DR   InterPro; IPR001275; DM_DNA-binding.
 DR   Pfam; PF00751; DM-domain; 1.
 FT   NON_TER       1      1
 FT   NON_TER      31     31
 SQ   SEQUENCE   31 AA;  3734 MW;  211519BE4019B92D CRC64;

    Q9DFI1  Length: 31  July 16, 2003 13:43  Type: P  Check: 8097  ..
 Found using 'var2' (kam816.key)

               |-----|
1       VVSALKGHKRYRRWKDCMCVKCTLIAERQRV
               9     15


                     -- Search Statistics --

Times:                    CPU                    Total Elapsed
                       00:00:00.00               00:00:00.00

Number of sequences searched:                        6
Number of sequence hits:                             6
Number of separate matches:                          6
Number of sequence hits saved:                       0
```

```
> O <
O|  |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "var3_pir" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  var3 (AA) ID   var3 AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         lf

Selected files:

        File : var3pir.pep

                    -- Output Parameters --

Format Options:                         File Options:
  Nucleic acid code matching   Exact     Indirect file                 No
  Find non-matching hits only  No        Sequence or key file          No
  Report key used              Yes       List of hits                  Yes
  Note position of hit         Yes       Hit display                   Yes
  Display full annotations     Yes       Name and annotations          Yes
  Sequence context             50


                    -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 c82922 ; TOIG of: c82922  check: 8625  from: 1  to: 521
  (from "var3pir.pep")
  TOIG of: c82922  check: 8625  from: 1  to: 521

 P1;C82922 - methionyl-tRNA synthetase UU197 [imported] - Ureaplasma urealyticum
 C;Species: Ureaplasma urealyticum
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Feb-2003
 C;Accession: C82922
 R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell,
  G.H.
 submitted to GenBank, February 2000
 A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views
  of a minimal genome and sexually transmitted pathogen.
 A;Reference number: A82870
 A;Accession: C82922
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-521 <GLA>
 A;Cross-references: GB:AE002119; GB:AF222894; NID:g6899153; PIDN:AAF30604.1;
  GSPDB:GN00123; UUSP:UU197
 A;Experimental source: serovar 3; biovar 1
 C;Genetics:
 A;Gene: metS; UU197
 A;Genetic code: SGC3
 C;Superfamily: methionyl-tRNA synthetase

    C82922  Length: 521  July 16, 2003 13:44  Type: P  Check: 8625  ..
 Found using 'var3' (kam816.key)

                                                    |-----|
1       MLKQKKFFISTPIYYSSGNPHIGHAYTTIIADVLARYKRLFGYDVFFLTGMDEHGQKIQQ
                                                   35    41

61      KAFEENISPKALVDRNSIIFLNLWKRLNISF

...

-------
1 match found in sequence:
 g64691 ; TOIG of: g64691  check: 6636  from: 1  to: 270
  (from "var3pir.pep")
  TOIG of: g64691  check: 6636  from: 1  to: 270

 P1;G64691 - UDP-N-acetylglucosamine acyltransferase - Helicobacter pylori
  (strain 26695)
 C;Species: Helicobacter pylori
 C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 18-Jun-1999
 C;Accession: G64691
 R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.;
  Fleischmann, R.D.; Ketchum, K.A.; Klenk, H.P.; Gill, S.; Dougherty, B.A.;
  Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus,
  B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.;
  Fitzegerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne,
  J.D.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman,
  J.M.; Fujii, C.; Bowman, C.; Watthey, L.
 Nature 388, 539-547, 1997
 A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.;
  Fraser, C.M.; Venter, J.C.
 A;Title: The complete genome sequence of the gastric pathogen Helicobacter
  pylori.
 A;Reference number: A64520; MUID:97394467; PMID:9252185
 A;Accession: G64691
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-270 <TOM>
 A;Cross-references: GB:AE000637; GB:AE000511; NID:g2314536; PIDN:AAD08418.1;
  PID:g2314545; TIGR:HP1375
 C;Superfamily: UDP-N-acetylglucosamine acyltransferase

    G64691  Length: 270  July 16, 2003 13:44  Type: P  Check: 6636  ..
 Found using 'var3' (kam816.key)

...

                                                                |-----|
162     AKGCMIAGKSALGKDVPPYCTVEGNRAFIRGLNRHRMRQLLESKDIDFIYALYKRLFRPI
                                                               212   218

222     PSLRESAKLELEEHANNPFVKEICSFILESSRGVAYKSSEYSSEEKQ

...

-------
1 match found in sequence:
 g71826 ; TOIG of: g71826  check: 5582  from: 1  to: 270
  (from "var3pir.pep")
  TOIG of: g71826  check: 5582  from: 1  to: 270

 P1;G71826 - udp-n-acetylglucosamine acyltransferase - Helicobacter pylori
  (strain J99)
 C;Species: Helicobacter pylori
 A;Variety: strain J99
 C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 18-Jun-1999
 C;Accession: G71826
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.;
  Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge, B.L.; Carmel, G.; Tummino, P.J.;
  Caruso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.;
  Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Trust, T.J.
 Nature 397, 176-180, 1999
 A;Title: Genomic sequence comparison of two unrelated isolates of the human
  gastric pathogen Helicobacter pylori.
 A;Reference number: A71800; MUID:99120557; PMID:9923682
```

```
A;Accession: G71826
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-270 <ARN>
A;Cross-references: GB:AE001551; GB:AE001439; NID:g4155887; PIDN:AAD06863.1;
 PID:g4155893
A;Experimental source: strain J99
C;Genetics:
A;Gene: lpxA
C;Superfamily: UDP-N-acetylglucosamine acyltransferase

   G71826  Length: 270  July 16, 2003 13:44  Type: P  Check: 5582  ..
Found using 'var3' (kam816.key)

...

                                                                |-----|
162     AKGCMIAGKSALGKDVPPYCTVEGNRAFIRGLNRHRMRQLLESKDIDFIHALYKRLFRPI
                                                              212   218

222     PSLRESAKLELEEHANNPFVKEICSFILESSRGVAYKSSEYSSEEKQ

...

-------
1 match found in sequence:
 h72742 ; TOIG of: h72742  check: 4517  from: 1  to: 172
  (from "var3pir.pep")
  TOIG of: h72742  check: 4517  from: 1  to: 172

 P1;H72742 - hypothetical protein APE0472 - Aeropyrum pernix (strain K1)
 C;Species: Aeropyrum pernix
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 08-Sep-2000
 C;Accession: H72742
 R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no,
  K.; Takahashi, M.; Sekine, M.; Baba, S.; Ankai, A.; Kosugi, H.; Hosoyama, A.;
  Fukui, S.; Nagai, Y.; Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.;
  Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.;
  Aoki, K.; Kubota, K.; Nakamura, Y.; Nomura, N.; Sako, Y.; Kikuchi, H.
 DNA Res. 6, 83-101, 1999
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic
  Crenarchaeon, Aeropyrum pernix K1.
 A;Reference number: A72450; MUID:99310339; PMID:10382966
 A;Accession: H72742
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-172 <KAW>
 A;Cross-references: DDBJ:AP000059; NID:g5103911; PIDN:BAA79436.1; PID:d1043222;
  PID:g5104120
 A;Experimental source: strain K1
 C;Genetics:
 A;Gene: APE0472
 C;Superfamily: Aeropyrum pernix hypothetical protein APE0472

    H72742  Length: 172  July 16, 2003 13:44  Type: P  Check: 4517  ..
 Found using 'var3' (kam816.key)

...

                                                          |-----|
114     IALSLKLEAPLWTEDKALLRASFKSNLYVALDTESVENLLHGEPLESIREKLYKKLFKQ
                                                        164   170

-------
1 match found in sequence:
 t27570 ; TOIG of: t27570  check: 7198  from: 1  to: 721
  (from "var3pir.pep")
  TOIG of: t27570  check: 7198  from: 1  to: 721

 P1;T27570 - hypothetical protein ZC434.6 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
 C;Accession: T27570
 R;Wilkinson, J.
 submitted to the EMBL Data Library, July 1996
 A;Reference number: Z20388
 A;Accession: T27570
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-721 <WIL>
 A;Cross-references: EMBL:Z75714; PIDN:CAB00063.1; GSPDB:GN00019; CESP:ZC434.6
 A;Experimental source: clone ZC434
 C;Genetics:
 A;Gene: CESP:ZC434.6
 A;Map position: 1
 A;Introns: 49/1; 244/2; 357/2; 440/2; 524/3; 611/3; 699/1
 C;Superfamily: Caenorhabditis elegans hypothetical protein ZC434.6

    T27570  Length: 721  July 16, 2003 13:44  Type: P  Check: 7198  ..
 Found using 'var3' (kam816.key)

...

                                                                |-----|
500     AILAAAADYVGVETDEVVAKVDKKLITTIFDCLITSNFWFDCDFMQKLDGGRYHKLFNSY
                                                              550   556

560     GFNQKSTYISMESHTAFPTVLHWLTIFALGSDKETLNVKSEKSCSHL

...

-------
1 match found in sequence:
 t41055 ; TOIG of: t41055  check: 5944  from: 1  to: 1315
  (from "var3pir.pep")
  TOIG of: t41055  check: 5944  from: 1  to: 1315

 P1;T41055 - membrane atpase - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C;Accession: T41055
 R;Hilbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1998
 A;Reference number: Z21967
 A;Accession: T41055
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1315 <HIL>
 A;Cross-references: EMBL:AL031324; PIDN:CAA20449.1; GSPDB:GN00068;
  SPDB:SPCC1672.11c
 A;Experimental source: strain 972h-; cosmid c1672
 C;Genetics:
 A;Gene: SPDB:SPCC1672.11c
 A;Map position: 3

   T41055  Length: 1315  July 16, 2003 13:44  Type: P  Check: 5944  ..
 Found using 'var3' (kam816.key)

...

                                                                |-----|
1251    SKSFQNFLIFAGVIYYLLAASGQNYIFISMTNFISHLNNRLLNRRTKVSKKLYKRLFADL
                                                              1301  1307

1311    QNEQV

-------
1 match found in sequence:
 t49926 ; TOIG of: t49926  check: 7827  from: 1  to: 397
  (from "var3pir.pep")
  TOIG of: t49926  check: 7827  from: 1  to: 397

 P1;T49926 - hypothetical protein F17I14.90 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Mar-2001
 C;Accession: T49926
```

R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.;
Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24490
A;Accession: T49926
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <BEV>
A;Cross-references: EMBL:AL353994; GSPDB:GN00063; ATSP:F17I14.90
A;Experimental source: cultivar Columbia; BAC clone F17I14
C;Genetics:
A;Gene: ATSP:F17I14.90
A;Map position: 5
A;Introns: 76/3; 86/3; 124/1; 164/3; 198/3; 315/3; 330/3; 369/3
C;Superfamily: Arabidopsis thaliana hypothetical protein F17J16.20

T49926 Length: 397 July 16, 2003 13:44 Type: P Check: 7827 ..
Found using 'var3' (kam816.key)

...

```
                                                              |-----|
339     SICVSVYSVVVGILGMNIPFPWNIKKHMFKWVVSGTATVCAILFVIIMSFARYKKLFGF
                                                             389   395
```

-- Search Statistics --

| Times: | CPU | Total Elapsed |
|---|---|---|
| | 00:00:00.00 | 00:00:00.00 |

| | |
|---|---|
| Number of sequences searched: | 7 |
| Number of sequence hits: | 7 |
| Number of separate matches: | 7 |
| Number of sequence hits saved: | 0 |

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "var3_sp" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  var3 (AA) ID   var3 AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         lf

Selected files:

        File : var3sp.pep

                    -- Output Parameters --

Format Options:                         File Options:
  Nucleic acid code matching   Exact     Indirect file                 No
  Find non-matching hits only  No        Sequence or key file          No
  Report key used              Yes       List of hits                  Yes
  Note position of hit         Yes       Hit display                   Yes
  Display full annotations     Yes       Name and annotations          Yes
  Sequence context             50


                    -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 atc9schpo ; Probable cation-transporting ATPase C1672.11c (EC 3.6.3.-).
  (from "var3sp.pep")
  TOIG of: atc9_schpo  check: 5944  from: 1  to: 1315

 ID   ATC9_SCHPO     STANDARD;      PRT;  1315 AA.
 AC   O74431;
 DT   30-MAY-2000 (Rel. 39, Created)
 DT   30-MAY-2000 (Rel. 39, Last sequence update)
 DT   15-SEP-2003 (Rel. 42, Last annotation update)
 DE   Probable cation-transporting ATPase C1672.11c (EC 3.6.3.-).
 GN   SPCC1672.11C.
 OS   Schizosaccharomyces pombe (Fission yeast).
 OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC   Schizosaccharomyces.
 OX   NCBI_TaxID=4896;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=972;
 RX   MEDLINE=21848401; PubMed=11859360;
 RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
 RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT   "The genome sequence of Schizosaccharomyces pombe.";
 RL   Nature 415:871-880(2002).
 CC   -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
 CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC   -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
 CC       ATPases). Subfamily V.
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; AL031324; CAA20449.1; -.
 DR   PIR; T41055; T41055.
 DR   GeneDB_SPombe; SPCC1672.11c; -.
 DR   InterPro; IPR001757; ATPase_E1-E2.
 DR   InterPro; IPR005834; Hydrolase.
 DR   InterPro; IPR006544; P-ATPase-V.
 DR   Pfam; PF00122; E1-E2_ATPase; 1.
 DR   Pfam; PF00702; Hydrolase; 1.
 DR   PRINTS; PR00119; CATATPASE.
 DR   TIGRFAMs; TIGR01494; ATPase_P-type; 5.
 DR   TIGRFAMs; TIGR01657; P-ATPase-V; 1.
 DR   PROSITE; PS00154; ATPASE_E1_E2; 1.
 KW   Hypothetical protein; Hydrolase; Transmembrane; Phosphorylation;
 KW   Magnesium; ATP-binding.
 FT   DOMAIN        1    139       CYTOPLASMIC (POTENTIAL).
 FT   TRANSMEM    140    161       POTENTIAL.
 FT   DOMAIN      162    167       EXTRACELLULAR (POTENTIAL).
 FT   TRANSMEM    168    190       POTENTIAL.
 FT   DOMAIN      191    324       CYTOPLASMIC (POTENTIAL).
 FT   TRANSMEM    325    347       POTENTIAL.
 FT   DOMAIN      348    350       EXTRACELLULAR (POTENTIAL).
 FT   TRANSMEM    351    369       POTENTIAL.
 FT   DOMAIN      370    531       CYTOPLASMIC (POTENTIAL).
 FT   TRANSMEM    532    551       POTENTIAL.
 FT   DOMAIN      552    564       EXTRACELLULAR (POTENTIAL).
 FT   TRANSMEM    565    586       POTENTIAL.
 FT   DOMAIN      587   1072       CYTOPLASMIC (POTENTIAL).
 FT   TRANSMEM   1073   1091       POTENTIAL.
 FT   DOMAIN     1092   1099       EXTRACELLULAR (POTENTIAL).
 FT   TRANSMEM   1100   1117       POTENTIAL.
 FT   DOMAIN     1118   1135       CYTOPLASMIC (POTENTIAL).
 FT   TRANSMEM   1136   1159       POTENTIAL.
 FT   DOMAIN     1160   1180       EXTRACELLULAR (POTENTIAL).
 FT   TRANSMEM   1181   1203       POTENTIAL.
 FT   DOMAIN     1204   1216       CYTOPLASMIC (POTENTIAL).
 FT   TRANSMEM   1217   1236       POTENTIAL.
 FT   DOMAIN     1237   1253       EXTRACELLULAR (POTENTIAL).
 FT   TRANSMEM   1254   1274       POTENTIAL.
 FT   DOMAIN     1275   1315       CYTOPLASMIC (POTENTIAL).
 FT   MOD_RES     619    619       PHOSPHORYLATION (BY SIMILARITY).
 FT   METAL      1015   1015       MAGNESIUM (BY SIMILARITY).
 FT   METAL      1019   1019       MAGNESIUM (BY SIMILARITY).
 SQ   SEQUENCE   1315 AA;  148786 MW;  081A550E59F2B0B5 CRC64;
```

```
ATC9_SCHPO  Length: 1315  July 16, 2003 13:45  Type: P  Check: 5944  ..
Found using 'var3' (kam816.key)

...

                                                              |-----|
1251    SKSFQNFLIFAGVIYYLLAASGQNYIFISMTNFISHLNNRLLNRRTKVSKKLYKRLFADL
                                                            1301  1307

1311    QNEQV

-------
1 match found in sequence:
 lpxahelpj ; Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransfe
  (from "var3sp.pep")
  TOIG of: lpxa_helpj  check: 5582  from: 1  to: 270

 ID   LPXA_HELPJ     STANDARD;      PRT;   270 AA.
 AC   Q9ZJL7;
 DT   16-OCT-2001 (Rel. 40, Created)
 DT   16-OCT-2001 (Rel. 40, Last sequence update)
 DT   28-FEB-2003 (Rel. 41, Last annotation update)
 DE   Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
 DE   (EC 2.3.1.129) (UDP-N-acetylglucosamine acyltransferase).
 GN   LPXA OR JHP1289.
 OS   Helicobacter pylori J99 (Campylobacter pylori J99).
 OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC   Helicobacteraceae; Helicobacter.
 OX   NCBI_TaxID=85963;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RX   MEDLINE=99120557; PubMed=9923682;
 RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA   Trust T.J.;
 RT   "Genomic sequence comparison of two unrelated isolates of the human
 RT   gastric pathogen Helicobacter pylori.";
 RL   Nature 397:176-180(1999).
 CC   -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF LIPID A, A
 CC       PHOSPHORYLATED GLYCOLIPID THAT ANCHORS THE LIPOPOLYSACCHARIDE TO
 CC       THE OUTER MEMBRANE OF THE CELL (BY SIMILARITY).
 CC   -!- CATALYTIC ACTIVITY: (R)-3-hydroxytetradecanoyl-[acyl-carrier
 CC       protein] + UDP-N-acetylglucosamine = [acyl-carrier protein] +
 CC       UDP-3-O-(3-hydroxytetradecanoyl)-N-acetylglucosamine.
 CC   -!- PATHWAY: Lipid A biosynthesis; first step.
 CC   -!- SUBUNIT: Homotrimer (By similarity).
 CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC   -!- SIMILARITY: BELONGS TO THE TRANSFERASE HEXAPEPTIDE REPEAT FAMILY.
 CC       LPXA SUBFAMILY.
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; AE001551; AAD06863.1; -.
 DR   PIR; G71826; G71826.
 DR   HSSP; P10440; 1LXA.
 DR   HAMAP; MF_00387; -; 1.
 DR   InterPro; IPR001451; Hexapep_transf.
 DR   Pfam; PF00132; hexapep; 7.
 DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; 2.
 KW   Transferase; Acyltransferase; Lipid A biosynthesis; Lipid synthesis;
 KW   Repeat; Complete proteome.
 SQ   SEQUENCE   270 AA;  29790 MW;  BC72E581E0226E23 CRC64;

  LPXA_HELPJ  Length: 270  July 16, 2003 13:45  Type: P  Check: 5582  ..
Found using 'var3' (kam816.key)

...

                                                              |-----|
162     AKGCMIAGKSALGKDVPPYCTVEGNRAFIRGLNRHRMRQLLESKDIDFIHALYKRLFRPI
                                                            212   218

222     PSLRESAKLELEEHANNPFVKEICSFILESSRGVAYKSSEYSSEEKQ

...

-------
1 match found in sequence:
 lpxahelpy ; Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransfe
  (from "var3sp.pep")
  TOIG of: lpxa_helpy  check: 6636  from: 1  to: 270

 ID   LPXA_HELPY     STANDARD;      PRT;   270 AA.
 AC   O25927;
 DT   16-OCT-2001 (Rel. 40, Created)
 DT   16-OCT-2001 (Rel. 40, Last sequence update)
 DT   28-FEB-2003 (Rel. 41, Last annotation update)
 DE   Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase
 DE   (EC 2.3.1.129) (UDP-N-acetylglucosamine acyltransferase).
 GN   LPXA OR HP1375.
 OS   Helicobacter pylori (Campylobacter pylori).
 OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC   Helicobacteraceae; Helicobacter.
 OX   NCBI_TaxID=210;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=26695 / ATCC 700392;
 RX   MEDLINE=97394467; PubMed=9252185;
 RA   Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA   Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 RA   Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA   Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
 RA   McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA   Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA   Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
 RA   Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA   Venter J.C.;
 RT   "The complete genome sequence of the gastric pathogen Helicobacter
 RT   pylori.";
 RL   Nature 388:539-547(1997).
 CC   -!- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF LIPID A, A
 CC       PHOSPHORYLATED GLYCOLIPID THAT ANCHORS THE LIPOPOLYSACCHARIDE TO
 CC       THE OUTER MEMBRANE OF THE CELL (BY SIMILARITY).
 CC   -!- CATALYTIC ACTIVITY: (R)-3-hydroxytetradecanoyl-[acyl-carrier
 CC       protein] + UDP-N-acetylglucosamine = [acyl-carrier protein] +
 CC       UDP-3-O-(3-hydroxytetradecanoyl)-N-acetylglucosamine.
 CC   -!- PATHWAY: Lipid A biosynthesis; first step.
 CC   -!- SUBUNIT: Homotrimer (By similarity).
 CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC   -!- SIMILARITY: BELONGS TO THE TRANSFERASE HEXAPEPTIDE REPEAT FAMILY.
 CC       LPXA SUBFAMILY.
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; AE000637; AAD08418.1; -.
 DR   PIR; G64691; G64691.
 DR   HSSP; P10440; 1LXA.
 DR   TIGR; HP1375; -.
 DR   HAMAP; MF_00387; -; 1.
 DR   InterPro; IPR001451; Hexapep_transf.
 DR   Pfam; PF00132; hexapep; 8.
```

```
DR   PROSITE; PS00101; HEXAPEP_TRANSFERASES; 1.
KW   Transferase; Acyltransferase; Lipid A biosynthesis; Lipid synthesis;
KW   Repeat; Complete proteome.
SQ   SEQUENCE   270 AA;  29855 MW;  3668DF2ED248BA4E CRC64;

 LPXA_HELPY  Length: 270  July 16, 2003 13:45  Type: P  Check: 6636  ..
Found using 'var3' (kam816.key)

...

                                                                |-----|
162     AKGCMIAGKSALGKDVPPYCTVEGNRAFIRGLNRHRMRQLLESKDIDFIYALYKRLFRPI
                                                              212   218

222     PSLRESAKLELEEHANNPFVKEICSFILESSRGVAYKSSEYSSEEKQ

...

-------
1 match found in sequence:
 nicacaeel ; Nicastrin homolog precursor.
  (from "var3sp.pep")
  TOIG of: nica_caeel  check: 7198  from: 1  to: 721

 ID   NICA_CAEEL     STANDARD;      PRT;   721 AA.
 AC   Q23316;
 DT   15-JUL-1998 (Rel. 36, Created)
 DT   15-JUL-1998 (Rel. 36, Last sequence update)
 DT   28-FEB-2003 (Rel. 41, Last annotation update)
 DE   Nicastrin homolog precursor.
 GN   APH-2 OR ZC434.6.
 OS   Caenorhabditis elegans.
 OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC   Rhabditidae; Peloderinae; Caenorhabditis.
 OX   NCBI_TaxID=6239;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=Bristol N2;
 RA   Wilkinson J.;
 RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
 RN   [2]
 RP   REVISIONS.
 RA   Jones S.J.M.;
 RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
 RN   [3]
 RP   POSSIBLE FUNCTION, AND GENE NAME.
 RX   MEDLINE=20445163; PubMed=10993067;
 RA   Yu G., Nishimura M., Arawaka S., Levitan D., Zhang L., Tandon A.,
 RA   Song Y.-Q., Rogaeva E., Chen F., Kawarai T., Supala A., Levesque L.,
 RA   Yu H., Yang D.-S., Holmes E., Milman P., Liang Y., Zhang D.M.,
 RA   Xu D.H., Sato C., Rogaev E., Smith M., Janus C., Zhang Y.,
 RA   Aebersold R., Farrer L.S., Sorbi S., Bruni A., Fraser P.E.,
 RA   St George-Hyslop P.H.;
 RT   "Nicastrin modulates presenilin-mediated notch/glp-1 signal
 RT   transduction and betaAPP processing.";
 RL   Nature 407:48-54(2000).
 CC   -!- FUNCTION: PLAYS A ROLE IN EMBRYONIC GLP-1 SIGNALING.
 CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC   -!- SIMILARITY: BELONGS TO THE NICASTRIN FAMILY.
 CC   ---------------------------------------------------------------------
 CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC   the European Bioinformatics Institute. There are no restrictions on its
 CC   use by non-profit institutions as long as its content is in no way
 CC   modified and this statement is not removed. Usage by and for commercial
 CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC   or send an email to license@isb-sib.ch).
 CC   ---------------------------------------------------------------------
 DR   EMBL; Z75714; CAB00063.1; -.
 DR   PIR; T27570; T27570.
 DR   WormPep; ZC434.6; CE15229.
 KW   Transmembrane; Glycoprotein; Signal.
 FT   SIGNAL        1     16       POTENTIAL.
 FT   CHAIN        17    721       NICASTRIN HOMOLOG.
 FT   DOMAIN       17    678       EXTRACELLULAR (POTENTIAL).
 FT   TRANSMEM    679    699       POTENTIAL.
 FT   DOMAIN      700    721       CYTOPLASMIC (POTENTIAL).
 FT   CARBOHYD     40     40       N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT   CARBOHYD    181    181       N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT   CARBOHYD    271    271       N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT   CARBOHYD    328    328       N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT   CARBOHYD    409    409       N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT   CARBOHYD    625    625       N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ   SEQUENCE   721 AA;  81383 MW;  904063F6E9CFB0D1 CRC64;

  NICA_CAEEL  Length: 721  July 16, 2003 13:45  Type: P  Check: 7198  ..
 Found using 'var3' (kam816.key)

 ...

                                                                |-----|
 500     AILAAAADYVGVETDEVVAKVDKKLITTIFDCLITSNFWFDCDFMQKLDGGRYHKLFNSY
                                                               550   556

 560     GFNQKSTYISMESHTAFPTVLHWLTIFALGSDKETLNVKSEKSCSHL

 ...

 -------
 1 match found in sequence:
  symurepa ; Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
   (from "var3sp.pep")
   TOIG of: sym_urepa  check: 8625  from: 1  to: 521

  ID   SYM_UREPA      STANDARD;      PRT;   521 AA.
  AC   Q9PQU6;
  DT   28-FEB-2003 (Rel. 41, Created)
  DT   28-FEB-2003 (Rel. 41, Last sequence update)
  DT   28-FEB-2003 (Rel. 41, Last annotation update)
  DE   Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
  DE   (MetRS).
  GN   METG OR METS OR UU197.
  OS   Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
  OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
  OX   NCBI_TaxID=134821;
  RN   [1]
  RP   SEQUENCE FROM N.A.
  RC   STRAIN=Serovar 3;
  RX   MEDLINE=20500219; PubMed=11048724;
  RA   Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
  RA   Cassell G.H.;
  RT   "The complete sequence of the mucosal pathogen Ureaplasma
  RT   urealyticum.";
  RL   Nature 407:757-762(2000).
  CC   -!- FUNCTION: Is required not only for elongation of protein synthesis
  CC       but also for the initiation of all mRNA translation through
  CC       initiator tRNA(fMet) aminoacylation (By similarity).
  CC   -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
  CC       diphosphate + L-methionyl-tRNA(Met).
  CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
  CC   -!- SUBUNIT: Monomer (By similarity).
  CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
  CC   -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
  CC       MetG subfamily 2A.
  CC   ---------------------------------------------------------------------
  CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC   the European Bioinformatics Institute. There are no restrictions on its
  CC   use by non-profit institutions as long as its content is in no way
  CC   modified and this statement is not removed. Usage by and for commercial
  CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
  CC   or send an email to license@isb-sib.ch).
  CC   ---------------------------------------------------------------------
  DR   EMBL; AE002119; AAF30604.1; -.
  DR   HSSP; P23395; 1A8H.
  DR   HAMAP; MF_01228; -; 1.
```

```
DR   InterPro; IPR002300; tRNA-synt_1a.
DR   InterPro; IPR001412; tRNA-synt_I.
DR   InterPro; IPR002304; tRNA-synt_met.
DR   Pfam; PF00133; tRNA-synt_1; 2.
DR   PRINTS; PR01041; TRNASYNTHMET.
DR   TIGRFAMs; TIGR00398; metG; 1.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
KW   Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW   Metal-binding; Zinc; Complete proteome.
FT   SITE         14     24       "HIGH" REGION.
FT   SITE        306    310       "KMSKS" REGION.
FT   METAL       129    129       ZINC (BY SIMILARITY).
FT   METAL       132    132       ZINC (BY SIMILARITY).
FT   METAL       151    151       ZINC (BY SIMILARITY).
FT   METAL       155    155       ZINC (BY SIMILARITY).
FT   BINDING     309    309       ATP (BY SIMILARITY).
SQ   SEQUENCE   521 AA;  60830 MW;  F9CD1A428B132DC5 CRC64;

 SYM_UREPA  Length: 521  July 16, 2003 13:45  Type: P  Check: 8625  ..
Found using 'var3' (kam816.key)

                                        |-----|
1       MLKQKKFFISTPIYYSSGNPHIGHAYTTIIADVLARYKRLFGYDVFFLTGMDEHGQKIQQ
                                       35    41

61      KAFEENISPKALVDRNSIIFLNLWKRLNISF

...


                    -- Search Statistics --

Times:                   CPU                  Total Elapsed
                    00:00:00.00               00:00:00.00

Number of sequences searched:                        5
Number of sequence hits:                             5
Number of separate matches:                          5
Number of sequence hits saved:                       0
```

```
> O <
O| |O  IntelliGenetics
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

                -- Outline of search "var3_spt" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.
Selected sequence key from "kam816.key":
  var3 (AA) ID   var3 AA preliminary pattern
1       followed by
2         a or k or g
2         l or r
2         y
2         k or r or h
2         k or r or h
2         lf

Selected files:

        File : var3spt.pep

                    -- Output Parameters --

Format Options:                        File Options:
  Nucleic acid code matching   Exact      Indirect file                 No
  Find non-matching hits only  No         Sequence or key file          No
  Report key used              Yes        List of hits                  Yes
  Note position of hit         Yes        Hit display                   Yes
  Display full annotations     Yes        Name and annotations          Yes
  Sequence context             50


                    -- Run Parameters --

  Run mode                     Batch
    Time to start comparison   now
    Notify at end of run       No

-------
1 match found in sequence:
 q8h1g7 ; MRS2-8.
  (from "var3spt.pep")
  TOIG of: q8h1g7  check: 8275  from: 1  to: 380

 ID   Q8H1G7      PRELIMINARY;      PRT;   380 AA.
 AC   Q8H1G7;
 DT   01-MAR-2003 (TrEMBLrel. 23, Created)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   MRS2-8.
 GN   MRS2-8.
 OS   Arabidopsis thaliana (Mouse-ear cress).
 OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX   NCBI_TaxID=3702;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RX   MEDLINE=21624993; PubMed=11752386;
 RA   Li L., Tutone A.F., Drummond R.S.M., Gardner R.C., Luan S.;
 RT   "A novel family of magnesium transport genes in Arabidopsis.";
 RL   Plant Cell 13:2761-2775(2001).
 RN   [2]
 RP   SEQUENCE FROM N.A.
 RA   Gardner R.C., Drummond R.S.M.;
 RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AY150293; AAN73218.1; -.
 SQ   SEQUENCE   380 AA;  42912 MW;  51E0E0EEF85840F8 CRC64;

    Q8H1G7  Length: 380  July 16, 2003 13:46  Type: P  Check: 8275  ..
 Found using 'var3' (kam816.key)

...

                                                                |-----|
322     SICVSVYSVVVGILGMNIPFPWNIKKHMFKWVVSGTATVCAILFVTIMSFARYKKLFGF
                                                              372   378

.
-------
1 match found in sequence:
 q8i4p8 ; Hypothetical protein.
  (from "var3spt.pep")
  TOIG of: q8i4p8  check: 9498  from: 1  to: 179

 ID   Q8I4P8      PRELIMINARY;      PRT;   179 AA.
 AC   Q8I4P8;
 DT   01-MAR-2003 (TrEMBLrel. 23, Created)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Hypothetical protein.
 GN   PFL2595W.
 OS   Plasmodium falciparum (isolate 3D7).
 OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX   NCBI_TaxID=36329;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=3D7;
 RX   MEDLINE=22255705; PubMed=12368864;
 RA   Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA   Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA   Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA   Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
 RA   Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
 RA   Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA   McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA   Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA   Fraser C.M., Barrell B.;
 RT   "Genome sequence of the human malaria parasite Plasmodium
 RT   falciparum.";
 RL   Nature 419:498-511(2002).
 DR   EMBL; AE014852; AAN36601.1; -.
 KW   Hypothetical protein.
 SQ   SEQUENCE   179 AA;  20824 MW;  58981BC75546B4C4 CRC64;

    Q8I4P8  Length: 179  July 16, 2003 13:46  Type: P  Check: 9498  ..
 Found using 'var3' (kam816.key)

...

                                                                |-----|
120     YNDLSKQFTLEELHTVLDNLEERPSNEDLYNIWNHVLGITKEGFEDKIFLALYRRLFKKI
                                                              170   176

180

-------
1 match found in sequence:
 q8ksr8 ; WaaJ.
  (from "var3spt.pep")
  TOIG of: q8ksr8  check: 4841  from: 1  to: 336

 ID   Q8KSR8      PRELIMINARY;      PRT;   336 AA.
 AC   Q8KSR8;
 DT   01-OCT-2002 (TrEMBLrel. 22, Created)
 DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   WaaJ.
 GN   WAAJ.
 OS   Salmonella enterica subsp. arizonae.
 OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC   Enterobacteriaceae; Salmonella.
 OX   NCBI_TaxID=59203;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SARC 5;
RA   Kaniuk N.A., Monteiro M.A., Parker C.T., Whitfield C.;
RT   "Molecular diversity of the genetic loci responsible for
RT   lipopolysaccharide core oligosaccharide assembly within the genus
RT   Salmonella.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF511116; AAM48163.1; -.
SQ   SEQUENCE   336 AA;  38891 MW;  87E86B135325BDA1 CRC64;

   Q8KSR8  Length: 336  July 16, 2003 13:46  Type: P  Check: 4841  ..
Found using 'var3' (kam816.key)

...

                                                                |-----|
257     NTILIHYTGATKPWHAWANYPSVIYYKNARLNSPWKDFPGKDARTIVEFKKRYKHLFVQG
                                                              307   313

317     HYFKGLLAGSAYLYRKVFHK

-------
1 match found in sequence:
 q8p6w5 ; Cytochrome B561.
  (from "var3spt.pep")
  TOIG of: q8p6w5  check: 2337  from: 1  to: 187

 ID   Q8P6W5      PRELIMINARY;      PRT;   187 AA.
 AC   Q8P6W5;
 DT   01-OCT-2002 (TrEMBLrel. 22, Created)
 DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Cytochrome B561.
 GN   YODB OR XCC2850.
 OS   Xanthomonas campestris (pv. campestris).
 OC   Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC   Xanthomonadaceae; Xanthomonas.
 OX   NCBI_TaxID=340;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=ATCC 33913 / NCPPB 528;
 RX   MEDLINE=22022145; PubMed=12024217;
 RA   da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA   Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA   Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA   Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA   Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA   Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA   Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA   Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA   Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA   Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA   Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA   Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA   Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA   Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA   Setubal J.C., Kitajima J.P.;
 RT   "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT   host specificities.";
 RL   Nature 417:459-463(2002).
 DR   EMBL; AE012398; AAM42122.1; -.
 DR   InterPro; IPR000516; Ni_hydr_CytB.
 DR   Pfam; PF01292; Ni_hydr_CYTB; 1.
 KW   Complete proteome.
 SQ   SEQUENCE   187 AA;  21077 MW;  423897ECBC7DD560 CRC64;

   Q8P6W5  Length: 187  July 16, 2003 13:46  Type: P  Check: 2337  ..
Found using 'var3' (kam816.key)

...

                                                          |-----|
```

```
112     TSGLRPFRWFGLFNVPKLSAPDEQLRAVSHTVHEWGFWLLIAVVLAHAGAALYHHLFQRD
                                                              162   168

172     ATLTRMLPRGWLSPKP

-------
1 match found in sequence:
 q8qne6 ; EsV-1-133.
  (from "var3spt.pep")
  TOIG of: q8qne6  check: 9763  from: 1  to: 353

 ID   Q8QNE6      PRELIMINARY;      PRT;   353 AA.
 AC   Q8QNE6;
 DT   01-JUN-2002 (TrEMBLrel. 21, Created)
 DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE   EsV-1-133.
 OS   Ectocarpus siliculosus virus.
 OC   Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
 OX   NCBI_TaxID=37665;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=EsV-1;
 RA   Delaroque N., Bothe G., Pohl T., Knippers R., Mueller D.G., Boland W.;
 RT   "The complete nucleotide sequence of the Ectocarpus siliculosus virus
 RT   genome.";
 RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
 DR   EMBL; AF204951; AAK14551.1; -.
 SQ   SEQUENCE   353 AA;  39951 MW;  87151FE1BD44D094 CRC64;

   Q8QNE6  Length: 353  July 16, 2003 13:46  Type: P  Check: 9763  ..
Found using 'var3' (kam816.key)

...

                                                          |-----|
76      DMPDAIDMIETANGVMISNIDADKMGARLNAMTSRWLEKRGRFVENDPEAGLYKKLFRLK
                                                              126   132

136     FIYNHLDEVYSDKYNIYQAAYVDPAYTVRAVLCFVMQIISFAALMKE

...

-------
1 match found in sequence:
 q8vvq9 ; ORF43.
  (from "var3spt.pep")
  TOIG of: q8vvq9  check: 3014  from: 1  to: 720

 ID   Q8VVQ9      PRELIMINARY;      PRT;   720 AA.
 AC   Q8VVQ9;
 DT   01-MAR-2002 (TrEMBLrel. 20, Created)
 DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   ORF43.
 OS   Staphylococcus aureus.
 OG   Plasmid ETB plasmid.
 OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX   NCBI_TaxID=1280;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=TY4;
 RX   MEDLINE=21562640; PubMed=11705958;
 RA   Yamaguchi T., Hayashi T., Takami H., Ohnisi M., Murata T.,
 RA   Nakayama K., Asakawa K., Ohara M., Komatsuzawa H., Sugai M.;
 RT   "Complete Nucleotide Sequence of a Staphylococcus aureus Exfoliative
 RT   Toxin B Plasmid and Identification of a Novel ADP-Ribosyltransferase,
 RT   EDIN-C.";
 RL   Infect. Immun. 69:7760-7771(2001).
 DR   EMBL; AP003088; BAB78441.1; -.
 DR   InterPro; IPR003593; AAA_ATPase.
 DR   InterPro; IPR001140; ABC_TM_transpt.
 DR   InterPro; IPR003439; ABC_transporter.
```

```
DR   InterPro; IPR005074; Peptidase_C39.
DR   Pfam; PF00664; ABC_membrane; 1.
DR   Pfam; PF00005; ABC_tran; 1.
DR   Pfam; PF03412; Peptidase_C39; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding; Plasmid.
SQ   SEQUENCE   720 AA;  81680 MW;  4B1A3AD33C3591CA CRC64;

   Q8VVQ9  Length: 720  July 16, 2003 13:46  Type: P  Check: 3014  ..
Found using 'var3' (kam816.key)

...

                                                              |-----|
654     HVYDELSRLKCTKIVIAHRFATVFNADKIVVMDHGEIVEVGAHNELLNKNGLYRKLFNAE
                                                            704   710

714     FKKECGV

-------
1 match found in sequence:
q9dwd8 ; PR52.
 (from "var3spt.pep")
 TOIG of: q9dwd8  check: 2812  from: 1  to: 530

ID   Q9DWD8      PRELIMINARY;      PRT;   530 AA.
AC   Q9DWD8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PR52.
GN   R52.
OS   Rat cytomegalovirus (strain Maastricht).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Betaherpesvirinae; Muromegalovirus.
OX   NCBI_TaxID=79700;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Maastricht;
RX   MEDLINE=20366325; PubMed=10906222;
RA   Vink C., Beuken E., Bruggeman C.A.;
RT   "Complete DNA sequence of the rat cytomegalovirus genome.";
RL   J. Virol. 74:7656-7665(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Maastricht;
RX   MEDLINE=20473137; PubMed=11018281;
RA   Gruijthuijsen Y.K., Beuken E., Bruggeman C.A., Vink C.;
RT   "Rat cytomegalovirus R89 is a highly conserved gene which expresses a
RT   spliced transcript.";
RL   Virus Res. 69:119-130(2000).
DR   EMBL; AF232689; AAF99151.1; -.
DR   InterPro; IPR002597; Herpes_env.
DR   Pfam; PF01673; Herpes_env; 1.
SQ   SEQUENCE   530 AA;  60538 MW;  1F9BAA2A685CA744 CRC64;

   Q9DWD8  Length: 530  July 16, 2003 13:46  Type: P  Check: 2812  ..
Found using 'var3' (kam816.key)

...

                                                              |-----|
375     LDLPADLKNKDLSSYLADAETPAKATPPGSRDPFELDAYTYLILRQVGVVGLYKHLFTDP
                                                            425   431

435     VCAANIRATDTSVLFFDVPNEYLNEVKLAICSTNAYPSAVERDFWLY

...

-------
1 match found in sequence:
q91xd5 ; Hypothetical 44.8 kDa protein.
 (from "var3spt.pep")
 TOIG of: q91xd5  check: 7827  from: 1  to: 397

ID   Q9LXD5      PRELIMINARY;      PRT;   397 AA.
AC   Q9LXD5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 44.8 kDa protein.
GN   F17I14_90.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA   Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL353994; CAB89358.1; -.
DR   InterPro; IPR002523; Mg2+_transptCorA.
DR   Pfam; PF01544; CorA; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   397 AA;  44752 MW;  496F484DD659E4A2 CRC64;

   Q9LXD5  Length: 397  July 16, 2003 13:46  Type: P  Check: 7827  ..
Found using 'var3' (kam816.key)

...

                                                              |-----|
339     SICVSVYSVVVGILGMNIPFPWNIKKHMFKWVVSGTATVCAILFVIIMSFARYKKLFGF
                                                            389   395

-------
1 match found in sequence:
q9s4d0 ; Transporter.
 (from "var3spt.pep")
 TOIG of: q9s4d0  check: 3300  from: 1  to: 720

ID   Q9S4D0      PRELIMINARY;      PRT;   720 AA.
AC   Q9S4D0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Transporter.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=1280;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C55;
RX   MEDLINE=99346225; PubMed=10417203;
RA   Navaratna M.A., Sahl H.G., Tagg J.R.;
RT   "Identification of genes encoding two-component lantibiotic production
RT   in Staphylococcus aureus C55 and other phage group II S. aureus
RT   strains and demonstration of an association with the exfoliative toxin
RT   B gene.";
RL   Infect. Immun. 67:4268-4271(1999).
DR   EMBL; AF147744; AAD47014.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR001140; ABC_TM_transpt.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR005074; Peptidase_C39.
DR   Pfam; PF00664; ABC_membrane; 1.
DR   Pfam; PF00005; ABC_tran; 1.
DR   Pfam; PF03412; Peptidase_C39; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
```

```
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding.
SQ   SEQUENCE   720 AA;  81668 MW;  EEDAC9DF33C13A24 CRC64;

   Q9S4D0  Length: 720  July 16, 2003 13:46  Type: P  Check: 3300  ..
Found using 'var3' (kam816.key)

...

                                                                |-----|
654     HVYDELSRLKCTKIVIAHRFATVFNADKIVVMDHGEIVEVGAHNELLNKNGLYRKLFNAE
                                                               704   710

714     FKKECGV

-------
1 match found in sequence:
 q9yew0 ; Hypothetical protein APE0472.
  (from "var3spt.pep")
  TOIG of: q9yew0  check: 4517  from: 1  to: 172

 ID   Q9YEW0      PRELIMINARY;      PRT;   172 AA.
 AC   Q9YEW0;
 DT   01-NOV-1999 (TrEMBLrel. 12, Created)
 DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE   Hypothetical protein APE0472.
 GN   APE0472.
 OS   Aeropyrum pernix.
 OC   Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC   Desulfurococcaceae; Aeropyrum.
 OX   NCBI_TaxID=56636;
 RN   [1]
 RP   SEQUENCE FROM N.A.
 RC   STRAIN=K1;
 RX   MEDLINE=99310339; PubMed=10382966;
 RA   Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA   Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT   "Complete genome sequence of an aerobic hyper-thermophilic
 RT   crenarchaeon, Aeropyrum pernix K1.";
 RL   DNA Res. 6:83-101(1999).
 DR   EMBL; AP000059; BAA79436.1; -.
 DR   InterPro; IPR002716; PIN.
 DR   InterPro; IPR006596; PINc.
 DR   InterPro; IPR000531; TonB_boxC.
 DR   Pfam; PF01850; PIN; 1.
 DR   SMART; SM00670; PINc; 1.
 DR   PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
 KW   Hypothetical protein; Complete proteome.
 SQ   SEQUENCE   172 AA;  19704 MW;  D735D4BC0D1A83FA CRC64;

   Q9YEW0  Length: 172  July 16, 2003 13:46  Type: P  Check: 4517  ..
Found using 'var3' (kam816.key)

...

                                                          |-----|
114     IALSLKLEAPLWTEDKALLRASFKSNLYVALDTESVENLLHGEPLESIREKLYKKLFKQ
                                                         164   170


                    -- Search Statistics --

Times:                  CPU                 Total Elapsed
                     00:00:00.00            00:00:00.00

Number of sequences searched:                    10
Number of sequence hits:                         10
Number of separate matches:                      10

Number of sequence hits saved:                   0
```